EXPEDITED PROCEDURE
EXAMINING GROUP 1653

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s):	Donoho *et al.*	Group Art Unit:	1653
Application No.:	09/784,358	Examiner:	S. Snedden
Filed:	02/15/2001	Atty. Docket No.:	LEX-0134-USA

Title: Novel Human Thrombospondin Repeat Proteins and Polynucleotides Encoding the Same

RESPONSE TO OFFICE ACTION DATED MAY 21, 2003

Mail Stop AF
Commissioner for Patents
Alexandria, VA 22313

Sir:

The Applicants acknowledge the receipt of the Office Action ("the Action") mailed on May 21, 2003 (Paper No. 12), which has been carefully reviewed and studied. Reexamination and reconsideration of the application is requested in view of the following remarks. In order to facilitate the Examiner's evaluation of the application, Applicants have attempted to address the rejections in Paper No. 12 in the same order in which they were originally raised.

The response is timely filed, and Applicants believe no fees are due in connection with this response. However, the Commissioner is authorized to charge any required fees or credit any overpayment to Deposit Account No. 50-0892.

RESPONSE

I. Status of the Claims

No claims have been cancelled. No claims have been amended. No new claims have been added.

Claims 1, 3, and 5-7 are therefore presently pending in the case. For the convenience of the Examiner, a clean copy of the pending claims is attached hereto as **Exhibit A**.

II. Rejection of Claims 1, 3 and 5-7 Under 35 U.S.C. § 101

The Action first rejects claims 1, 3 and 5-7 under 35 U.S.C. § 101, as allegedly lacking a patentable utility. Applicants respectfully traverse.

Applicants pointed out in the response filed on February 13, 2003 ("the previous response") to the First Office Action in this case, which was issued on August 27, 2002 ("the First Action"), a sequence sharing nearly 100% percent identity at the protein level over an extended region of the claimed sequence is present in the leading scientific repository for biological sequence data (GenBank), and has been annotated by third party scientists who are *wholly unaffiliated with Applicants* as "ADAMTSL3" (GenBank accession number AF237652; alignment and GenBank report shown in **Exhibit B**). The Examiner questioned this assertion of utility because "'nearly 100% identity over and extend region' (*sic*) does not necessarily provide adequate support" for this assertion (Action at page 4, emphasis in original). The Examiner further stated that "Vernet *et al.* (US 20030059768) teach a protein that shares 99.9% identity with the sequence of SEQ ID NO:2 over the entire sequence" and that this "protein is described as (*sic*) Vernet *et al.* as a STE20-like trypsin inhibitor" (Action bridging pages 4 and 5, emphasis in original). There are a number a flaws in the Examiner's argument. First, while it is true that the alignment between SEQ ID NO:2 and the sequence disclosed in GenBank accession number AF237652 only covers 766 of the 1691 amino acids of SEQ ID NO:2 (with 765, or 99.9%, identity), the AF237652 sequence is clearly annotated in GenBank as a "partial cds" (see **Exhibit B**) and only contains 766 amino acids. However, examination of the manuscript referenced in the GenBank report for AF237652, Hirohata *et al.* (which, although still indicated as "unpublished", in fact was published in J. Biol. Chem. *277*:12182-12189, 2002; "Hirohata"; **Exhibit C**), clearly indicates that the full length ADAMTSL3 sequence is "1690 amino acids" in length (Hirohata at page 12187, second column). Thus, although the information in GenBank accession number AF237652 has not been updated to reflect the full length ADAMTSL3 sequence (or information about the Hirohata publication), the skilled artisan would readily recognize that Applicants' claimed sequence is, in fact, a full length ADAMTSL3 sequence.

Second, Applicants are completely at a loss to understand why the Examiner seems to be accepting the assertion of utility set forth in U.S. Patent Application 20030059768 at face value ("(t)he protein is described as (*sic*) Vernet *et al.* as a STE20-like trypsin inhibitor" (Action bridging pages 4 and 5)), while not accepting Applicants' assertion that the presently claimed sequence encodes an ADAMTS-like protein, even though third party scientists who are *wholly unaffiliated with Applicants*

have **confirmed Applicants assertion**, as described above. Applicants respectfully remind the Examiner that the legal test for utility simply involves an assessment of whether those skilled in the art would find any of the utilities described for the invention to be credible or believable. Given this GenBank annotation and the Hirohata manuscript, there can be no question that those skilled in the art would clearly believe that Applicants' sequence is an ADAMTSL3 sequence. The specification as originally filed details that the present sequence is associated with the extracellular matrix (at least at page 1, line 27), which is confirmed by Hirohata (see **Exhibit C**). Given the well established biological relevance of ADAMTS-like proteins, those of skill in the art would readily appreciate the utility of the present sequence in numerous applications, as described below. Thus, the Examiner's argument fails to support the assertion that the present claims lack a patentable utility.

As set forth in the previous response, as just one example of the utility of the present nucleotide sequences, the specification details on page 5, lines 2-4, that the present nucleotide sequences have utility in assessing gene expression patterns using high-throughput DNA chips. Such "DNA chips" clearly have utility, as evidenced by hundreds of issued U.S. Patents, as exemplified by U.S. Patent Nos. 5,445,934, 5,556,752, 5,744,305, 5,837,832, 6,156,501 and 6,261,776. As the present sequences are specific markers of the human genome (see below), and such specific markers are targets for the discovery of drugs that are associated with human disease, those of skill in the art would instantly recognize that the present nucleotide sequences would be an ideal, novel candidate for assessing gene expression using such DNA chips. Given the widespread utility of such "gene chip" methods using *public domain* gene sequence information, there can be little doubt that the use of the presently described *novel* sequences would have great utility in such DNA chip applications. Clearly, compositions that enhance the utility of such DNA chips, such as the presently claimed nucleotide sequences, must in themselves be useful.

Evidence of the "real world" substantial utility of the present invention is further provided by the fact that there is an entire industry established based on the use of gene sequences or fragments thereof in a gene chip format. Perhaps the most notable gene chip company is Affymetrix. However, there are many companies which have, at one time or another, concentrated on the use of gene sequences or fragments, in gene chip and non-gene chip formats, for example: Gene Logic, ABI-Perkin-Elmer, HySeq and Incyte. In addition, one such company (Rosetta Inpharmatics) was viewed to have such "real world" value that it was acquired by large a pharmaceutical company (Merck) for significant sums of money (net equity value of the transaction was $620 million). The "real world" substantial industrial

utility of gene sequences or fragments would, therefore, appear to be widespread and well established. Clearly, persons of skill in the art, as well as venture capitalists and investors, readily recognize the utility, both scientific and commercial, of genomic data in general, and specifically human genomic data. Billions of dollars have been invested in the human genome project, resulting in useful genomic data (see, *e.g.*, Venter *et al.*, 2001, Science *291*:1304). The results have been a stunning success as the utility of human genomic data has been widely recognized as a great gift to humanity (see, *e.g.*, Jasny and Kennedy, 2001, Science *291*:1153). Clearly, the usefulness of human genomic data, such as the presently claimed nucleic acid molecules, is substantial and credible (worthy of billions of dollars and the creation of numerous companies focused on such information) and well-established (the utility of human genomic information has been clearly understood for many years).

The Examiner questions this asserted utility, stating that such a use does "not constitute a specific or substantial utility" (Action at page 4). The Examiner seems to be confusing the requirements of a **specific** utility with a **unique** utility. As clearly set forth by the Federal Circuit in *Carl Zeiss Stiftung v. Renishaw PLC*, 20 USPQ2d 1101 (Fed. Cir. 1991):

> An invention need not be the best or only way to accomplish a certain result, and it need only be useful to some extent and in certain applications: "[T]he fact that an invention has only limited utility and is only operable in certain applications is not grounds for finding a lack of utility." *Envirotech Corp. v. Al George, Inc.*, 221 USPQ 473, 480 (Fed. Cir. 1984)

The fact that other nucleotide sequences can be used to track gene expression does not mean that the use of Applicants' sequence to track gene expression is not a specific utility. If every invention were required to have a unique utility, the Patent and Trademark Office would no longer be issuing patents on batteries, automobile tires, golf balls, golf clubs, and treatments for a variety of human diseases, such as cancer, just to name a few particular examples, because the utility of each of these compositions is applicable to the broad class in which each of these compositions falls: all batteries have the same utility, specifically to provide electrical power; all automobile tires have the same utility, specifically for use on automobiles; all golf balls and golf clubs have the same utility, specifically for use in the game of golf; and all cancer treatments have the same utility, specifically, to treat cancer. However, only the briefest perusal of virtually any issue of the Official Gazette provides numerous examples of patents being granted on each of the above compositions nearly every week. Furthermore, if a composition needed to be unique to be patented, the entire class and subclass system would be an effort in futility, as the class and subclass system serves solely to group such common inventions, which would not be required if each invention needed to have a unique utility. Thus, the present sequence clearly meets the

requirements of 35 U.S.C. § 101.

Although Applicants need only make one credible assertion of utility to meet the requirements of 35 U.S.C. § 101 (*Raytheon v. Roper*, 220 USPQ 592 (Fed. Cir. 1983); *In re Gottlieb*, 140 USPQ 665 (CCPA 1964); *In re Malachowski*, 189 USPQ 432 (CCPA 1976); *Hoffman v. Klaus*, 9 USPQ2d 1657 (Bd. Pat. App. & Inter. 1988)), Applicants pointed out in the previous response that as a further example of the utility of the presently claimed polynucleotide, as described in the specification at least at page 9, line 37, the present nucleotide sequence has a specific utility in determining the genomic structure of the corresponding human chromosome, for example mapping the protein encoding regions. This is evidenced by the fact that SEQ ID NO: 1 can be used to map the 29 coding exons on chromosome 15 (present within four overlapping chromosome 15 clones; GenBank Accession Numbers AC027807, AC022684, AC116157 and AC087738; alignments and first page from the GenBank reports provided in **Exhibit D**). Clearly, the present polynucleotide provides exquisite specificity in localizing the specific region of human chromosome 15 that contains the gene encoding the given polynucleotide, a utility not shared by virtually any other nucleic acid sequences. In fact, it is this specificity that makes this particular sequence so useful. Early gene mapping techniques relied on methods such as Giemsa staining to identify regions of chromosomes. However, such techniques produced genetic maps with a resolution of only 5 to 10 megabases, far too low to be of much help in identifying specific genes involved in disease. The skilled artisan readily appreciates the significant benefit afforded by markers that map a specific locus of the human genome, such as the present nucleic acid sequence.

The Action also questions this utility, again stating that such a use does "not constitute a specific or substantial utility" (Action at page 4). First, Applicants respectfully remind the Examiner that only a minor percentage (2-4%) of the genome actually encodes exons, which in-turn encode amino acid sequences. Equally significant is that the claimed polynucleotide sequence defines how the encoded exons are actually spliced together to produce an active transcript (*i.e.*, the described sequences are useful for functionally defining exon splice-junctions). It is well-known that exon splice junctions can often be hot spots for erroneous events leading to cancer. The claimed sequences identify biologically verified exon splice junctions, as opposed to splice junctions that may have been bioinformatically predicted from genomic sequence alone. The specification also details that "sequences derived from regions adjacent to the intron/exon boundaries of the human gene can be used to design primers for use in amplification assays to detect mutations within the exons, introns, splice sites (*e.g.*, splice acceptor

and/or donor sites), *etc.*, that can be used in diagnostics and pharmacogenomics" (specification at page 10, lines 1-7). Applicants respectfully submit that the practical scientific value of biologically validated, expressed, spliced, and polyadenylated mRNA sequences is readily apparent to those skilled in the relevant biological and biochemical arts. Second, the Examiner again seems to be confusing the requirements of a **specific** utility with a **unique** utility. The fact that other nucleotide sequences can be used to identify exon splice junctions and map human chromosome 15 does not mean that these uses of Applicants' sequence are not specific utilities (*Carl Zeiss Stiftung v. Renishaw PLC, supra*). Thus, the present sequence clearly meets the requirements of 35 U.S.C. § 101.

It is important to note that it has been clearly established that a statement of utility in a specification must be accepted absent reasons why one skilled in the art would have reason to doubt the objective truth of such statement. *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA, 1974; "*Langer*"); *In re Marzocchi*, 439 F.2d 220, 224, 169 USPQ 367, 370 (CCPA, 1971). As clearly set forth in *Langer*:

> As a matter of Patent Office practice, a specification which contains a disclosure of utility which corresponds in scope to the subject matter sought to be patented must be taken as sufficient to satisfy the utility requirement of § 101 for the entire claimed subject matter unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope.

Langer at 297, emphasis in original. As set forth in the MPEP, "Office personnel must provide evidence sufficient to show that the statement of asserted utility would be considered 'false' by a person of ordinary skill in the art" (MPEP, Eighth Edition at 2100-40, emphasis added). Absent such evidence from the Examiner, the present claims clearly meet the requirements of 35 U.S.C. § 101.

The Action goes on to suggest that the claimed sequences lack utility because further research would be required in certain aspects of the invention. Even if, *arguendo*, further research might be required in certain aspects of the present invention, this does not preclude a finding that the invention has utility, as set forth by the Federal Circuit's holding in *In re Brana*, (34 USPQ2d 1436 (Fed. Cir. 1995), "*Brana*"), which clearly states that "pharmaceutical inventions, necessarily includes the expectation of further research and development" (*Brana* at 1442-1443, emphasis added). In assessing the question of whether undue experimentation would be required in order to practice the claimed invention, the key term is "undue", not "experimentation". *In re Angstadt and Griffin*, 190 USPQ 214 (CCPA 1976). The need for some experimentation does not render the claimed invention unpatentable. Indeed, a considerable amount of experimentation may be permissible if such experimentation is routinely practiced in the art. *In re Angstadt and Griffin, supra*; *Amgen, Inc. v.*

Chugai Pharmaceutical Co., Ltd., 18 USPQ2d 1016 (Fed. Cir. 1991). As a matter of law, it is well settled that a patent need not disclose what is well known in the art. *In re Wands*, 8 USPQ 2d 1400 (Fed. Cir. 1988).

Finally, the requirements set forth in the Action for compliance with 35 U.S.C. § 101 do not comply with the requirements set forth by the Patent and Trademark Office ("the PTO") itself for compliance with 35 U.S.C. § 101. While Applicants are well aware of the new Utility Guidelines set forth by the USPTO, Applicants respectfully point out that the current rules and regulations regarding the examination of patent applications is and always has been the patent laws as set forth in 35 U.S.C. and the patent rules as set forth in 37 C.F.R., not the Manual of Patent Examination Procedure or particular guidelines for patent examination set forth by the USPTO. Furthermore, it is the job of the judiciary, not the USPTO, to interpret these laws and rules. Applicants are unaware of any significant recent changes in either 35 U.S.C. § 101, or in the interpretation of 35 U.S.C. § 101 by the Supreme Court or the Federal Circuit that is in keeping with the new Utility Guidelines set forth by the USPTO. This is underscored by numerous patents that have been issued over the years that claim nucleic acid fragments that do not comply with the new Utility Guidelines. As examples of such issued U.S. Patents, the Examiner is invited to review U.S. Patent Nos. 5,817,479, 5,654,173, and 5,552,281 (each of which claims short polynucleotides), and recently issued U.S. Patent No. 6,340,583 (which includes no working examples), none of which contain examples of the "real-world" utilities that the Examiner seems to be requiring. As issued U.S. Patents are presumed to meet all of the requirements for patentability, including 35 U.S.C. §§ 101 and 112, first paragraph (see Section III, below), Applicants submit that the present polynucleotides must also meet the requirements of 35 U.S.C. § 101. While Applicants understand that each application is examined on its own merits, Applicants are unaware of any changes to 35 U.S.C. § 101, or in the interpretation of 35 U.S.C. § 101 by the Supreme Court or the Federal Circuit, since the issuance of these patents that render the subject matter claimed in these patents, which is similar to the subject matter in question in the present application, as suddenly non-statutory or failing to meet the requirements of 35 U.S.C. § 101. Thus, holding Applicants to a different standard of utility would be arbitrary and capricious, and, like other clear violations of due process, cannot stand.

For each of the foregoing reasons, as well as the reasons set forth in the previous response, Applicants submit that as the presently claimed nucleic acid molecules have been shown to have a substantial, specific, credible and well-established utility, the rejection of claims 1, 3 and 5-7 under

35 U.S.C. § 101 has been overcome, and request that the rejection be withdrawn.

III. Rejection of Claims 1, 3 and 5-7 Under 35 U.S.C. § 112, First Paragraph

The Action next rejects claims 1, 3 and 5-7 under 35 U.S.C. § 112, first paragraph, since allegedly one skilled in the art would not know how to use the invention, as the invention allegedly is not supported by a specific, substantial, and credible utility or a well-established utility. Applicants respectfully traverse.

Applicants submit that as claims 1, 3 and 5-7 have been shown to have "a specific, substantial, and credible utility", as detailed in section II above, the present rejection of claims 1, 3 and 5-7 under 35 U.S.C. § 112, first paragraph, cannot stand.

Applicants therefore request that the rejection of claims 1, 3 and 5-7 under 35 U.S.C. § 112, first paragraph, be withdrawn.

IV. Conclusion

The present document is a full and complete response to the Action. In conclusion, Applicants submit that, in light of the foregoing remarks, the present case is in condition for allowance, and such favorable action is respectfully requested. Should Examiner Snedden have any questions or comments, or believe that certain amendments of the claims might serve to improve their clarity, a telephone call to the undersigned Applicants' representative is earnestly solicited.

Respectfully submitted,

August 21, 2003
Date

David W. Hibler Reg. No. 41,071
Agent for Applicants

LEXICON GENETICS INCORPORATED
8800 Technology Forest Place
The Woodlands, TX 77381
(281) 863-3399

Exhibit A

Clean Version of The Pending Claims in U.S. Patent Application Ser. No. 09/784,358

1. (Previously Presented) An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1.

3. (Original) An isolated nucleic acid molecule comprising a nucleotide sequence that encodes the amino acid sequence shown in SEQ ID NO:2.

5. (Previously Presented) A recombinant expression vector comprising the isolated nucleic acid molecule of claim 3.

6. (Previously Presented) The recombinant expression vector of claim 5, wherein the isolated nucleic acid molecule comprises the nucleotide sequence of SEQ ID NO:1.

7. (Previously Presented) A host cell comprising the recombinant expression vector of claim 5.

```
>AF237652 ACCESSION:AF237652 NID: gi 13183077 gb AF237652.1 AF237652
          Homo sapiens a disintegrin-like and metalloprotease
          domain with thrombospondin type I motifs-like 3
          (ADAMTSL3) mRNA, partial cds
         Length = 2523

Identities = 765/766 (99%), Positives = 765/766 (99%)

Query: 1    MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 60
             MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY
Sbjct: 225  MASWTSPWWVLIGMVFMHSPLPQTTAEKSPGAYFLPEFALSPQGSFLEDTTGEQFLTYRY 404

Query: 61   DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 120
             DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN
Sbjct: 405  DDQTSRNTRSDEDKDGNWDAWGDWSDCSRTCGGGASYSLRRCLTGRNCEGQNIRYKTCSN 584

Query: 121  HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 180
             HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV
Sbjct: 585  HDCPPDAEDFRAQQCSAYNDVQYQGHYYEWLPRYNDPAAPCALKCHAQGQNLVVELAPKV 764

Query: 181  LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 240
             LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP
Sbjct: 765  LDGTRCNTDSLDMCISGICQAVGCDRQLGSNAKEDNCGVCAGDGSTCRLVRGQSKSHVSP 944

Query: 241  EKREENVIAVPLGSRSVRITVKGPAHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 300
             EKREENVIAVPLGSRSVRITVKGP HLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR
Sbjct: 945  EKREENVIAVPLGSRSVRITVKGPVHLFIESKTLQGSKGEHSFNSPGVFVVENTTVEFQR 1124

Query: 301  GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 360
             GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ
Sbjct: 1125GSERQTFKIPGPLMADFIFKTRYTAAKDSVVQFFFYQPISHQWRQTDFFPCTVTCGGGYQ 1304

Query: 361  LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 420
             LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP
Sbjct: 1305LNSAECVDIRLKRVVPDHYCHYYPENVKPKPKLKECSMDPCPSSDGFKEIMPYDHFQPLP 1484

Query: 421  RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 480
             RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP
Sbjct: 1485RWEHNPWTACSVSCGGGIQRRSFVCVEESMHGEILQVEEWKCMYAPKPKVMQTCNLFDCP 1664

Query: 481  KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 540
             KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK
Sbjct: 1665KWIAMEWSQCTVTCGRGLRYRVVLCINHRGEHVGGCNPQLKLHIKEECVIPIPCYKPKEK 1844

Query: 541  SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 600
             SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT
Sbjct: 1845SPVEAKLPWLKQAQELEETRIATEEPTFIPEPWSACSTTCGPGVQVREVKCRVLLTFTQT 2024

Query: 601  ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 660
             ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC
Sbjct: 2025ETELPEEECEGPKLPTERPCLLEACDESPASRELDIPLPEDSETTYDWEYAGFTPCTATC 2204

Query: 661  LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 720
             LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG
Sbjct: 2205LGGHQEAIAVCLHIQTQQTVNDSLCDMVHRPPAMSQACNTEPCPPRWHVGSWGPCSATCG 2384

Query: 721  VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIE 766
             VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIE
Sbjct: 2385VGIQTRDVYCLHPGETPAPPEECRDEKPHALQACNQFDCPPGWHIE 2522
```

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 2002 by The American Society for Biochemistry and Molecular Biology, Inc.
Vol. 277, No. 14, Issue of April 5, pp. 12182–12189, 2002
Printed in U.S.A.

Punctin, a Novel ADAMTS-like Molecule, ADAMTSL-1, in Extracellular Matrix*

Received for publication, October 5, 2001, and in revised form, January 17, 2002
Published, JBC Papers in Press, January 22, 2002, DOI 10.1074/jbc.M109665200

Satoshi Hirohata‡§, Lauren W. Wang‡§, Masaru Miyagi¶, Lin Yan¶, Michael F. Seldin‖, Douglas R. Keene, John W. Crabb¶, and Suneel S. Apte‡ ‡‡§§**

*From the ‡Department of Biomedical Engineering, Lerner Research Institute, the ‡‡Department of Orthopaedic Surgery, Cleveland Clinic Foundation, the ¶Department of Ophthalmic Research, Cole Eye Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195, the ‖Rowe Program in Genetics, Departments of Biological Chemistry and Medicine, University of California at Davis, Davis, California 956161, and the **Shriner's Hospital for Children, Portland, Oregon 97201*

Punctin (ADAMTSL-1) is a secreted molecule resembling members of the ADAMTS family of proteases. Punctin lacks the pro-metalloprotease and the disintegrin-like domain typical of this family but contains other ADAMTS domains in precise order including four thrombospondin type I repeats. Punctin is the product of a distinct gene on human chromosome 9p21-22 and mouse chromosome 4 that is expressed in adult skeletal muscle. His-tagged punctin expressed in stably transfected High-Five™ insect cells was purified to apparent homogeneity by Ni-chromatography of conditioned medium. The NH_2 terminus is not blocked and has the sequence EEDRD and so forth as determined by Edman degradation, demonstrating signal peptidase processing. Recombinant epitope-tagged punctin has a calculated mass of 59,991 Da but exhibits major molecular species of 61970 ± 6 Da and 62131 ± 5 Da as measured by liquid chromatography electrospray mass spectrometry. Punctin is a glycoprotein based on carbohydrate staining and liquid chromatography electrospray mass spectrometry glycopeptide analysis. Glycosylation occurs at a single *N*-linked site as demonstrated by altered electrophoretic migration of punctin expressed in the presence of tunicamycin A. Punctin contains disulfide bonds based on antibody accessibility and electrophoretic migration under reducing *versus* nonreducing conditions. Rotary shadowing demonstrates that punctin is hatchet-shaped having a globular region attached to a short stem. In transfected COS-1 cells, punctin is deposited in the cell substratum in a punctate fashion and is excluded from focal contacts. Punctin is the first member of a novel family of ADAMTS-like proteins that may have important functions in the extracellular matrix.

Metalloproteases responsible for extracellular (ECM)[1] turnover have a modular structure. Matrix metalloproteinases (MMPs) (1), a disintegrin-like and metalloprotease (ADAMs) (2), and proteases of the ADAMTS family (3, 4) are composed of characteristic domains arranged in a precise order that is the hallmark of each family. These enzymes are structurally and functionally bipartite consisting of an enzymatic domain attached to nonenzymatic or ancillary domains. The ancillary domains localize these proteases to substrates, the cell surface, or to the ECM. The ancillary domains of the gelatinases MMP-2 and MMP-9 are among the best studied of the substrate-binding domains. The fibronectin type II domains of the gelatinases are involved in binding to gelatin and some collagens as well as to fibronectin and heparin as in the case of MMP-2 (5, 6). The gelatin-binding domain of MMP-2 binds the matricellular proteins thrombospondin-1 (TSP1) and TSP2 (7). Although neither is a substrate for MMP-2, the interaction may mediate the clearance of MMP-2 and affect cell-adhesive properties (8). The MMP-2 hemopexin domain interacts with the carboxyl terminus of the tissue inhibitor of metalloproteases-2, facilitating pro-MMP-2 activation by membrane-type MMPs (1, 5, 6, 9). The MMP-2 hemopexin domain also interacts with a chemokine called monocyte chemoattractant protein-3, which allows its processing by the catalytic domain (10). The disintegrin domains of ADAMs such as ADAM-15 are implicated in cell-cell adhesion (2, 11, 12), and the ancillary domains of ADAMTS-1 are required for its binding to the ECM (13). In some ADAMs, the zinc-binding active site is nonfunctional, suggesting that they do not function as proteases at all but may instead have a primary role in adhesion via their ancillary domains (2).

With this background, it is conceptually possible that gene products containing only the ancillary domains of ADAMTS may have specific functions in cell-cell or cell-matrix interactions or may regulate ADAMTS proteases. We have identified an ADAMTS-like (ADAMTSL) molecule named punctin.[2]

* This work was supported in part by the Cleveland Clinic Foundation (to S. S. A.), a Yamanouchi USA Foundation Award (to S. S. A.), and National Institutes of Health Grants EY06603 (to J. W. C.) and HG00734 (to M. F. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF176313.

§ Both authors contributed equally to this work.

§§ To whom correspondence should be addressed: Dept. of Biomedical Engineering (ND20), Lerner Research Inst., Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland, OH 44195. Tel.: 216-445-3278; Fax: 216-445-4383; E-mail: aptes@bme.ri.ccf.org or www.lerner.ccf.org/pi/apte.html.

[1] The abbreviations used are: ECM, extracellular matrix; ADAMTSL, a disintegrin-like and metalloprotease domain with thrombospondin type I motifs like; ADAMTS, a disintegrin-like and metalloprotease domain with thrombospondin type I motifs; ADAM, a disintegrin-like and metalloprotease; MS, mass spectrometry; EST, expressed sequence tag; LC-ESMS, liquid chromatography-electrospray mass spectrometry; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; MMP, matrix metalloprotease; ORF, open reading frame; PBS, phosphate-buffered saline; RACE, rapid amplification of cDNA ends; TSP, thrombospondin; TS, thrombospondin type I domain; HexNAc, *N*-acetylhexosamine; NeuAc, *N*-acetylneuraminic acid.

[2] Approved gene symbols *ADAMTSL1* and ***Adamtsl1*** indicate human and mouse orthologs, respectively. The corresponding protein product of these genes, ADAMTSL-1, is designated by the trivial name punctin because of its punctate distribution beneath transfected cells.

which is the product of a gene distinct from any in the ADAMTS family and is composed of ADAMTS ancillary domains alone. We have purified and characterized recombinant punctin produced in insect cells, visualized it by electron microscopy, and demonstrated that it is a glycoprotein and a component of the ECM

EXPERIMENTAL PROCEDURES

cDNA Cloning and Sequence Analysis—Using BLAST programs from the National Center for Biotechnology Information, we scanned the data base of ESTs using the protein sequences of ADAMTS proteases previously cloned by us (4, 14) and identified a human EST (GenBank™ accession number AA482392 encoded by IMAGE clone 752797). The EST predicted a polypeptide with a similarity to the carboxyl half of cognate ADAMTS members but with no identities in GenBank™ or other protein and nucleotide data bases.

Using nested oligonucleotide primers based on the sequences at the 5′ and 3′ ends of the IMAGE clone insert and human skeletal muscle cDNA (Marathon cDNA, CLONTECH, Palo Alto, CA) as the template, we performed RACE and extended the cDNA at 5′ and 3′ ends by PCR essentially as described previously (4, 14).

Northern Blot Analysis—Multiple tissue Northern blots from adult human and mouse tissues (CLONTECH, Palo Alto, CA) were hybridized to a [α-^{32}P]dCTP-labeled punctin probe, a 1200-bp cDNA fragment from the 5′ end of the punctin coding sequence, followed by autoradiographic exposure for 7 days.

Chromosomal Mapping and Genomic Arrangement—To determine the chromosomal location of ***Adamtsl1***, we analyzed a panel of DNA samples from an interspecific cross that has been characterized for over 1200 genetic markers throughout the mouse genome (15). Markers can be seen on the worldwide web (www.informatics.jax.org/searches/cross-data_form.shtml) by entering "DNA Mapping Panel Data Sets" from the mouse genome data base and then selecting the "Seldin cross" and "Chromosome." Initially, DNA from the two parental mice, (C3H/HeJ-*gld*) and (C3H/HeJ-*gld* × *Mus spretus*) F_1), were digested with various restriction endonucleases and hybridized with the ***Adamtsl1*** cDNA probe (IMAGE clone 2076907 with GenBank™ accession number AI787975) to determine restriction fragment length variants for haplotype analyses. Gene linkage was determined by segregation analysis. Gene order was determined by analyzing all haplotypes and minimizing crossover frequency among all genes that were determined to be within a linkage group. This method resulted in the determination of the most probable gene order. To define the locus for *ADAMTSL1*, the human punctin cDNA sequence was used for BLAST searches of the human genome (Celera Sciences, Rockville, MD).

Generation and Characterization of Anti-punctin Antisera—The peptide (NH_2)-[C]YYPENIKPKPKLQE-(OH) located in the third TS domain of punctin (Fig. 1*B*) was synthesized using Fmoc (*N*-(9-fluorenyl) methoxycarbonyl) chemistry, purified by reverse-phase high-pressure liquid chromatography, and molecular weight was confirmed by MS (Alpha Diagnostic International, San Antonio, TX). A cysteine ([C]) residue was included at the NH_2 terminus for coupling to keyhole limpet hemocyanin. Peptide-keyhole limpet hemocyanin conjugate was dialyzed in PBS and used for immunization. Two New Zealand White male rabbits (7–8 pounds) were immunized with the conjugate (~200 μg/injection/rabbit, multiple intramuscular and subcutaneous sites) at biweekly intervals for 8 weeks. After an initial injection in Freund's complete adjuvant, subsequent injections were given in incomplete adjuvant. Antibody titer was measured by enzyme-linked immunosorbent assay using free peptide.

Immune sera were tested by Western blot analysis of extracts from COS-1 cells transiently transfected with punctin cDNA (see below). Although antisera from both rabbits (antisera 4112 and 4113) gave qualitatively similar results, the best signal/noise ratio was obtained with antiserum 4113. Affinity-purified antibodies were prepared by column chromatography of antiserum 4113 using the immobilized peptide immunogen.

Expression and Purification of Recombinant Punctin from Insect Cells—High-Five™ cells (Invitrogen) were routinely cultured on tissue culture plastic and maintained at 27 °C in Ultimate™ serum-free insect cell medium (Invitrogen) as per manufacturer's directions. The full-length punctin ORF was excised from pcDNA3.1/Myc-His B-TSL1 (see below) with *Eco*RI and *Not*I and ligated into the corresponding sites in pIZT/V5-His (Invitrogen). The resulting insect cell expression plasmid pIZT/V5-His-TSL1 generated punctin with a COOH-terminal V5 epitope and 6× His tag. pIZT/V5-His-TSL1 was transfected into High-Five™ cells using Insectin-Plus liposomes (Invitrogen) and plated onto 100-mm Petri dishes. After 48 h, antibiotic selection (500 μg/ml Zeocin, Invitrogen) was started and continued for 21 days. Colonies that survived selection were picked manually, expanded, and maintained in medium containing Zeocin (50 μg/ml). Punctin production by isolated colonies was tested by Western blot analysis of conditioned medium using anti-His monoclonal antibody (Invitrogen) and antibody 4113.

For protein production, cells were grown in suspension in either Ultimate™ serum-free insect cell medium or Express-Five serum-free medium containing heparin (5 units/ml, Invitrogen). Production cultures were in spinner flasks, and culture medium was stored at −80 °C with 1 mM phenylmethylsulfonyl fluoride until use. For purification, medium was dialyzed into binding buffer (20 mM sodium phosphate, 500 mM NaCl, pH 7.8) containing 0.03% Brij-35 (Sigma). Purification was performed using 1-liter batches of dialyzed medium and a 5-ml Ni-Sepharose column (ProBond™, Invitrogen) on an fast protein liquid chromatography instrument (Bio-Rad, Hercules, CA). Following binding, the column was washed with three column volumes of binding buffer. A gradient of 0–42.5 mM imidazole in binding buffer was used to remove nonspecifically bound molecules from the column. Elution was with four column volumes of 250 mM imidazole in binding buffer, pH 7.0, containing 0.03% Brij-35. Elution was monitored by in-line UV and conductivity measurements. 2-ml fractions of eluate were collected and tested by Western blot analysis as described above. Fractions containing punctin were pooled. Protein concentration was determined using the Bradford assay (Bio-Rad) and by phenylthiocarbamyl amino acid analysis using an Applied Biosystems model 420H/130/920 automated analysis system (16).

Characterization of Recombinant Punctin—The NH_2-terminal sequence of recombinant punctin was determined by Edman degradation. Recombinant punctin (5 μg) was electrophoresed on 10% SDS-PAGE, electrotransferred to polyvinylidene difluoride membrane, and lightly stained with modified Coomassie Blue (Simply Blue Safe Stain, Invitrogen). Protein bands were excised and subjected to Edman degradation on an Applied Biosystems Procise 492 sequencer in the Molecular Biotechnology Core Facility of the Lerner Research Institute.

To probe for glycosylation, recombinant punctin (4 μg) was electrophoresed on 10% SDS-PAGE and stained for carbohydrate using a periodic acid-Schiff reaction-based method (Pro-Q fuchsia glycoprotein staining kit, Molecular Probes, Eugene, OR). In this reaction, Candy-Cane™ glycoprotein molecular weight standards consisting of alternate bands of glycosylated and unglycosylated proteins were used as controls. Glycoprotein staining was also performed after enzymatic deglycosylation of punctin with peptide *N*-glycosidase F. Deglycosylation of denatured as well as native punctin was performed with a commercially available kit (Bio-Rad) using bovine fetuin as a control. To investigate further whether *N*-linked carbohydrates were present in punctin, stably transfected insect cells were cultured in the presence or absence of tunicamycin A1 homolog (0.1 μg/ml culture medium, Sigma). Equal amounts of total protein from culture medium of tunicamycin-treated and untreated cells were assayed by Western blot with antibody 4113 at various time points after the addition of tunicamycin

Mass Spectrometry—The molecular mass of punctin was measured by MALDI-TOF and by LC-ESMS. MALDI-TOF was performed with a PerkinElmer Biosystems Voyager DE Pro-mass spectrometer using sinapinic acid as the matrix and bovine serum albumin as a calibration standard protein (17). MALDI-TOF MS measurements of intact punctin and naturally observed limited proteolysis fragments are reported ± 50% peak width (in Da) at half-maximal peak height. LC-ESMS was performed with a PerkinElmer Sciex API 3000 triple quadruple mass spectrometer (17, 18). Nitrogen was used as the nebulization gas at 40 p.s.i., and curtain gas was supplied from a nitrogen generator (Whatman model 75-72). For LC-ESMS of intact punctin, a scan range of 700–1800 *m/z* was used with 0.2 atomic mass unit steps, a scan time of 7.5 s, and at an orifice potential of 80 and 5000 V ion spray. Reverse phase-high-pressure liquid chromatography was done at a flow rate of 5 μl/min on a 5-μm Vydac C18 capillary column (0.3 × 150 mm, LC Packing) using an Applied Biosystems Model 140D high-pressure liquid chromatography system and aqueous acetonitrile/trifluoroacetic acid solvents with 100% of the eluant going to the mass spectrometer. ESMS measurements of intact punctin are reported as the mean ± S.E. (in Da).

For glycopeptide characterization, punctin was excised from a SDS-polyacrylamide gel (~1 μg/lane × 6 lanes), in-gel reduced with 10 mM dithiothreitol, cysteine-alkylated with 20 mM iodoacetamide in 400 mM ammonium bicarbonate, and digested with 0.2 μg of trypsin (Promega) overnight at 37 °C in 100 mM ammonium bicarbonate. Peptides from the in-gel tryptic digests were extracted with 60% acetonitrile containing 0.1% trifluoroacetic acid, dried in a Speed Vac, redissolved in 50 μl

of 0.1% trifluoroacetic acid, and analyzed by LC-ESMS using selective ion monitoring with the PE Sciex API 3000 triple quadruple mass spectrometer system as described above for intact protein analyses. Glycopeptides were selectively detected based on diagnostic sugar oxonium ions HexNAc + Hex (*m/z* 366) and *N*-acetylneuraminic acid (NeuAc) (*m/z* 292) (17). Carbohydrate marker ions at *m/z* 366 and 292 (dwell time 200 ms each) were monitored in a positive ion mode at a high orifice potential (180 V), whereas full scans at *m/z* 300–2300 (0.2 atomic mass unit steps, scan time 3.5 s) were acquired at a lower orifice potential (70 V). This way both intact parent ions and abundant marker ions were observed in the same *m/z* scan.

Rotary Shadowing and Electron Microscopy of Recombinant Punctin—Rotary shadowing was done essentially as described previously (19). A 30-μl sample of punctin at 100 μg/ml was mixed with 70 μl of glycerol and nebulized onto freshly cleaved mica using an airbrush. The sample was dried in a vacuum, and rotary shadowed using a platinum-carbon electron beam gun angled at 6° relative to the mica surface within a Balzers BAE 250 evaporator. The replica was backed with carbon, floated onto distilled water, and picked up onto 600 mesh grids. Photomicrographs were taken using a Philips 410 electron microscope operated at 80 kV.

Transient Expression of Tagged and Untagged Punctin in COS-1 Cells—An internal *Sac*I site and a flanking *Not*I site were used to remove a 1.5-kb fragment of IMAGE clone 752797 and ligate it into corresponding sites in IMAGE clone 2150669 corresponding to the 5′ end of the punctin cDNA to generate a complete ORF. *Eco*RI and *Not*I sites flanking this ORF were used to excise and clone the full-length coding sequence into pcDNA3.1/Myc-His (+) A (Invitrogen) for the expression of untagged punctin. To make constructs in which the *ADAMTSL1* ORF was in-frame with a carboxyl-terminal FLAG tag or a tandem *myc* tag and 6× His tag, PCR was performed with Advantage 2 polymerase (CLONTECH, Palo Alto, CA) using the full-length coding sequence as a template. The amplicons were cloned into the vectors pFLAG-CMV5c (Sigma) and pcDNA3.1/Myc-His B (Invitrogen) for expression with either a COOH-terminal FLAG tag or a COOH-terminal tandem *myc* tag and 6× His tag, respectively.

COS-1 cells (ATCC number CRL-1650) were grown on tissue culture plastic in Dulbecco's modified Eagle's medium:F-12 (1:1) (Lerner Research Institute Media Services) supplemented with 10% fetal bovine serum (Invitrogen) and antibiotics (100 units/ml of penicillin and 50 μg/ml streptomycin). 10^5 cells between passages 3 and 10 were transfected with untagged, FLAG-tagged, or *myc* + 6× His-tagged punctin using FuGENE 6 (Roche Molecular Biochemicals) as per manufacturer's recommendations, and cells were grown for an additional 24–48 h in serum-supplemented or serum-free medium. As a control, cells were transfected with the respective vector alone without insert. The medium was collected and concentrated 10-fold. Cells were harvested after detachment with 10 mM EDTA for 10–15 min at 37 °C. A complete detachment of cells was confirmed by phase-contrast microscopy. Fifty microliters of 2× Laemmli sample buffer was added to the wells, and the ECM was scraped off. Samples of cell lysate, medium, and ECM were separately electrophoresed under reducing conditions (samples were boiled following the addition of 10% (v/v) 2-mercaptoethanol) on 12% SDS-polyacrylamide gels and transferred to enhanced chemiluminescence (ECL)-Hybond (Amersham Biosciences, Inc.). Western blotting was performed using either anti-FLAG M2 antibody (diluted 1:500, Sigma), anti-His (COOH-terminal) antibody (diluted 1:1000, Invitrogen) or antibody 4113 (diluted 1:300) depending on the construct used for transfection. Antibody binding was detected using the appropriate peroxidase-labeled second antibody followed by ECL using reagents from Amersham Biosciences, Inc.

For immunocytochemistry, COS-1 cells were grown on glass coverslips in 35-mm diameter wells (in 6-well plates) and transiently transfected as described above in serum-supplemented or serum-free medium. The medium was removed 48 h after transfections. The cells were washed three times on ice with cold PBS containing 1 mM $CaCl_2$ and 1 mM $MgCl_2$ and incubated for 1 h on ice with 1 ml of culture medium containing anti-FLAG M2 monoclonal antibody (diluted 1:300, Sigma) or anti-punctin rabbit antisera (diluted 1:100) with gentle shaking. Cells were washed four times for 3 min each with cold PBS, fixed in 4% paraformaldehyde (w/v in PBS) (Sigma) on ice for 30 min with gentle shaking and then washed three times with PBS at ambient temperature. To quench free aldehyde groups, cells were treated with 75 mM ammonium chloride, 20 mM glycine for 10 min at ambient temperature, washed with PBS, and then blocked with 0.05% Triton X-100, 2% normal goat serum in PBS (10 min at ambient temperature). Finally, sections were incubated with the species-appropriate Texas Red-labeled goat secondary antibody (Jackson ImmunoResearch Laboratories, West Grove, PA) prior to coverslip mounting in Vectashield containing 4′,6-diamidino-2-phenylindole (Vector Laboratories, Inc., Burlingame, CA). The following control-immunostaining experiments were performed. COS-1 cells transfected with the vector alone or untransfected COS-1 cells were stained with the above antibodies, or transfected cells were stained with preimmune serum from the rabbits in which the polyclonal antibodies were produced.

To co-stain punctin and the actin cytoskeleton, cells were stained with anti-FLAG or anti-punctin antibodies as described above with the exception that the secondary antibodies included incubation with Alexa 488-phalloidin at recommended dilutions (Molecular Probes). In double immunostaining experiments following the immunolocalization of FLAG or punctin as described above, cells were permeabilized with 0.1% Triton X-100 in PBS for 20 min prior to staining with (*a*) monoclonal antibody to vinculin (1:100 dilution, Sigma) in combination with antiserum 4113 for the detection of punctin or (*b*) polyclonal antibody to focal adhesion kinase (1:200 dilution, Upstate Biotechnology, Lake Placid, NY) in combination with anti-FLAG monoclonal antibody M2 (Sigma) for the detection of punctin. A Texas Red-labeled antibody (Jackson ImmunoResearch Laboratories) was used for the detection of punctin, and Alexa 488-conjugated antibody (Molecular Probes) was used for the detection of vinculin or focal adhesion kinase.

RESULTS

Cloning of Punctin cDNA—We identified a novel EST (GenBank™ accession number AA482392) derived from pooled human melanocyte, fetal heart, and pregnant uterus with homology to ADAMTS proteases. The 1.5-kb insert of the corresponding IMAGE clone 752797 contained a long ORF encoding an amino-terminal TS domain, a cysteine-rich domain, a cysteine-free spacer domain, and three tandem TS modules followed by a short acidic peptide and stop codon (Fig. 1*a*). The stop codon and 3′-untranslated sequence were independently confirmed by 3′-RACE (clone pSHTSL1s3, Fig. 1*a*) as well as by another EST (GenBank™ accession number W47029). The 3′-untranslated region encoded in IMAGE clone 752797 contained a consensus polyadenylation signal (AATTAAA) followed by a poly(A) tail 14 nucleotides downstream. Completion of the full-length coding sequences by 5′-RACE predicted a putative signal peptide upstream of the central TS domain. The signal peptide was preceded by a methionine codon within a satisfactory Kozak consensus sequence (A at −3, G at +4 relative to ATG) (20) although there was no upstream in-frame stop codon. The 5′ sequence obtained by RACE was subsequently validated by independently cloned human and mouse ESTs (Genbank™ accession numbers A1459225 for human EST and AK020115 for mouse EST). The continuity of the cDNA clones was confirmed by PCR amplification of the full-length punctin ORF from human skeletal muscle cDNA (see below) as well as by identification of the encoding exons arranged sequentially on human chromosome 9 (Celera Genomics, Rockville, MD).

Primary Structure of Punctin Predicts an ADAMTS-like Protein—The predicted full-length punctin protein contains 525 amino acids and has the typical domain structure of the ancillary noncatalytic regions of an ADAMTS protease (Fig. 1*a*). The mature secreted form of punctin is 497 amino acids with a molecular mass of 55,240 Da and a calculated pI of 6.2. Like the ADAMTS proteases, each domain in punctin has an even number of cysteine residues. This observation suggests that each domain may have internal disulfide bonds (17 such bonds are predicted in punctin), and that punctin consists of a series of independently-folded and disulfide-bonded domains. Punctin contains no other domains apart from those described previously in the ADAMTS family. The punctin sequence contains one motif for *N*-linked glycosylation (21) at Asn^{223} (-Asn-*X*-Ser/Thr-, where *X* is any amino acid except Pro) and also contains a total of 75 Thr and Ser residues, where *O*-linked glycosylation might occur. (Fig. 1*b*).

The overall punctin sequence is most similar to human AD-

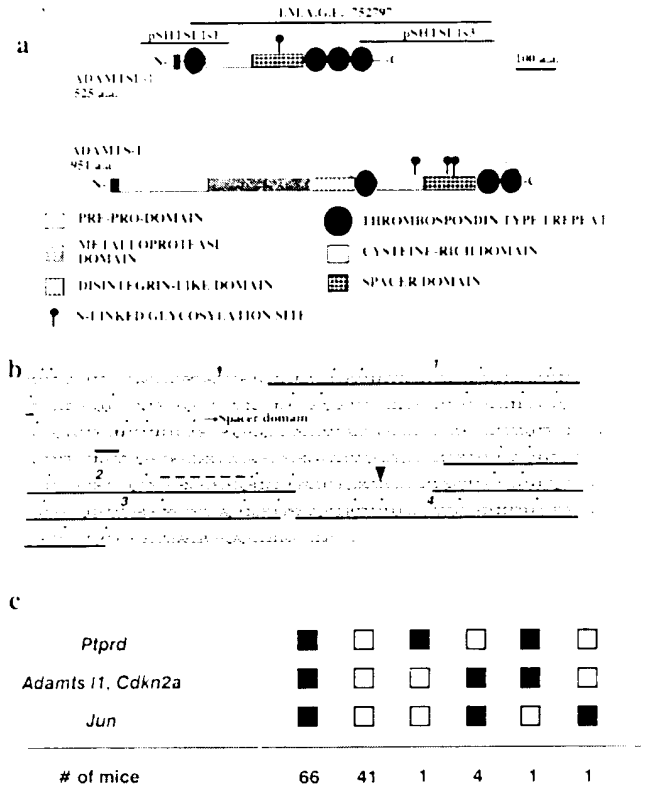

FIG. 1. *a*, domain organization of punctin/ADAMTSL-1 shown relative to ADAMTS-1, the prototypic ADAMTS. The cloning strategy used for determination of the complete primary structure is shown. The location of each cDNA clone relative to the protein domains indicates the regions it encodes. The key to the domains is shown at the *bottom* of the figure. *b*, the predicted amino acid sequence of punctin is shown using the single-letter amino acid code. TS modules are *underlined* with the *thick line* and are numbered sequentially from amino to carboxyl terminus. A consensus sequence for *N*-linked glycosylation is *overlined*. Cysteine residues are indicated by *asterisks*. The start of the spacer domain is indicated, the region between the NH_2-terminal TS domain and the spacer domain is the cysteine-rich domain. The *dashed line* indicates the peptide used for the generation of antibodies. The *arrow* indicates the signal peptidase cleavage site. The *arrowhead* indicates a putative proteolytic processing site between TS domains 2 and 3. *c*, segregation of ***Adamtsl1*** on mouse chromosome 4 in ((C3H/HeJ-*gld* × *M. spretus*) F_1 × C3H/HeJ-*gld*) interspecific backcross mice. *Filled boxes* represent the homozygous C3H pattern, and *open boxes* represent the F_1 pattern. The mapping of the reference loci in this interspecific cross has been previously described (15).

AMTSL-3 (68% identity, see below). Of the ADAMTS enzymes published to date, punctin is most similar to human ADAMTS-10 (35% identity). The punctin TS domains have a higher degree of similarity to other ADAMTS-like proteins and ADAMTS proteases than to TSP1 and TSP2. The greatest similarities, as indicated by percentage of identity of amino acid sequences identified by BLAST searches of the first TS domain of punctin to TS domains from various molecules, are as follows: human ADAMTSL-3, 80%; human ADAMTS-1, ADAMTS-6, and ADAMTS-10, 50%; mouse papilin, 47%; human ADAMTS-8, 44%; human ADAMTS-5, 42%; human TSP2, 40%; human TSP1, 38%. Like most TS domains in the ADAMTS family, punctin TS domains do not contain linear peptide sequences found in TSP1 that have been defined as heparin or CD-36 binding sequences, (22). They do not contain degenerate GAG binding sequences such as BBXB, where B is the basic amino acid and *X* is any amino acid (22).

Genomic Location of the Mouse and Human Punctin Genes and Tissue-specific Expression—The mapping of ***Adamtsl1*** in an interspecific cross resulted in the following most probable gene order (mean ± S.D.): ***Ptprd***-4.4 ± 2.0 centimorgan-***Adamtsl1, Cdkn2a***-1.8 ± 1.2 centimorgan-***Jun*** and placed ***Adamtsl1*** at a consensus position of 42.6 centimorgan on mouse chromosome 4 (Fig. 1*c*) in the vicinity of the interferon gene cluster. A search of the mouse genome data base (www.informatics.jax.org) did not reveal any pertinent genetic disorders near this locus.

The human-mouse homology maps (www3.ncbi.nlm.nih.gov/Omim/Homology/, accessed September 26, 2001) predict that the *ADAMTSL1* locus is on human chromosome 9p21-22. The predicted locus was confirmed by the analysis of the human genome sequence. The punctin ORF is encoded by 13 exons spanning >250 kb of genomic DNA mapping to 9p21.2-22.1. A search of the Online Mendelian Inheritance in Man site (www3.ncbi.nlm.nih.gov/Omim/) revealed three unsolved human disorders in the vicinity of the *ADAMTSL1* locus. Diaphyseal medullary stenosis with malignant fibrous histiocytoma (MIM112250) is linked to 9p22-p21, Friedreich's ataxia 2 (MIM601992) is linked to 9p23-p11, and neuropathy, distal hereditary motor, Jerash type (MIM605726) are linked to 9p21.1-p12.

ADAMTSL1 is primarily expressed in human and mouse skeletal muscle with a major message size of ~7.0 kb in both species (Fig. 2). A minor messenger RNA species of ~1.0 kb was also seen in some human tissues (Fig. 2, skeletal muscle, heart, colon, kidney, and liver). Expression was not detected in brain, colon, thymus, spleen, placenta, small intestine, lung, testis, ovary, or peripheral blood leukocytes.

Expression and Characterization of Recombinant Punctin—Punctin expressed in High-Five™ cells with tandem COOH-terminal V5 and 6× His epitopes was secreted into the conditioned medium of adherent as well as suspension cultures. Punctin was detected by antibody 4113 and anti-epitope tag antibodies as a ~60-kDa band under reducing conditions. It was substantially purified from the culture medium using Ni-chromatography (Fig. 3*a*). The purification scheme yielded a maximum of 200 μg/liter purified protein as determined by amino acid analysis. Electrophoresis and Western blotting of concentrated punctin preparations frequently demonstrated additional bands of molecular mass (~120 and ~180 kDa, data not shown), suggesting the formation of dimers and trimers at high concentrations.

The conformation of punctin appears to be maintained by disulfide bonds as evidenced by more rapid migration in SDS-PAGE under nonreducing conditions than under reducing conditions (Fig. 3*b*). Furthermore, on Western blots under nonreducing conditions, the protein was not detectable with antibody 4113 (data not shown), suggesting that the peptide epitope was not accessible without reduction of disulfide bonds. A mass analysis of His-tagged punctin by MALDI-TOF MS yielded a broad peak suggesting that the 60-kDa gel band contained major molecular species of 61,935 ± 595 and 60,873 ± 295 Da, respectively. LC-ESMS analyses of the intact protein defined more precisely the major molecular species to be 61,970 ± 6 and 62,131 ± 5, which are, respectively, 1979 and 2140 Da larger than the calculated mass (59,991) of tagged punctin based on amino acid sequence. NH_2-terminal sequencing of the polyvinylidene difluoride-immobilized 60-kDa protein revealed a single sequence, which commenced at Glu^{29} (*i.e.* Glu-Glu-Asp-Arg-Asp-Gly and so on).

Recombinant Punctin Is Glycosylated—Two closely spaced punctin bands were resolved by Western blot analysis of conditioned medium or purified protein, although Coomassie Blue staining of purified punctin always demonstrated a single band (Fig. 3*a*). A periodic acid-Schiff-based method of staining carbohydrate chains suggested that recombinant punctin is a gly-

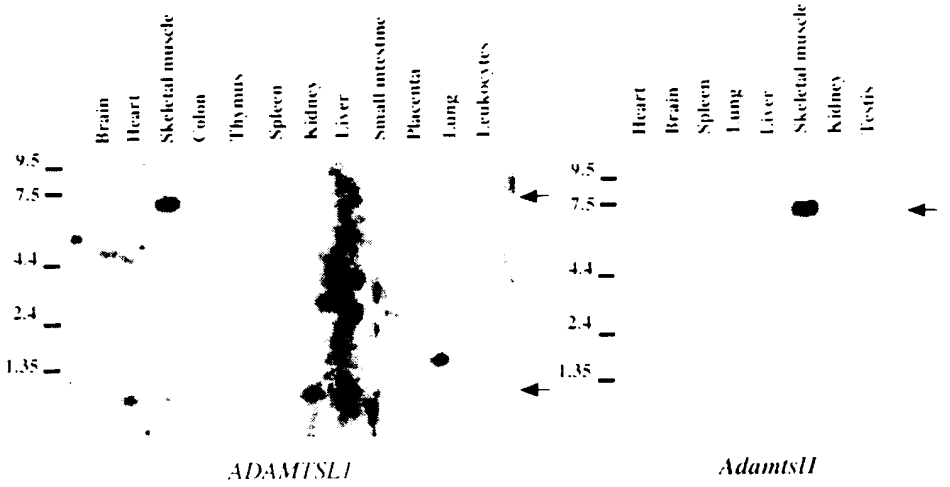

FIG. 2. **Northern analysis of expression of *ADAMTSL1* (*left*) and *Adamtsl1* (*right*) in adult human and mouse tissues, respectively.** Kilobase markers of RNA are shown at the *left* of each autoradiogram, and tissue origin is indicated above each lane. Hybridizing transcripts are indicated by *arrows*.

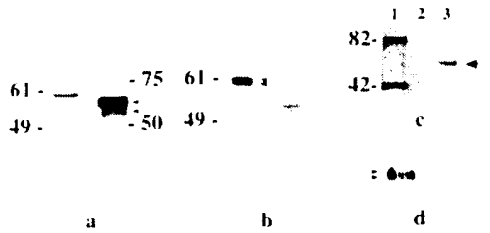

FIG. 3. **Analysis of epitope-tagged punctin purified by Ni-chromatography from insect cell culture medium.** *a*, Coomassie Blue (Simply Blue Safe Stain) staining of purified recombinant punctin on reducing SDS-PAGE (*left lane*) and Western blot analysis with anti-punctin antibody 4113 (*right lane*). *b*, Western blot analysis using anti-His tag monoclonal antibody on reducing (*left lane*) and nonreducing SDS-PAGE (*right lane*). *c*, glycoprotein staining of recombinant punctin (*lane 2* contains 0.6 μg, and *lane 3* contains 3 μg) using the periodic acid-Schiff procedure. Glycosylated CandyCane™ markers (1 μg/band) stained similarly are in *lane 1*. The *arrow* indicates stained punctin. *d*, Western analysis of culture medium from insect cell cultures treated without (*left lane*) or with (*right lane*) tunicamycin A for 72 h. Each lane contains 2.8 μg of total protein. *Double arrowheads* are used to indicate two molecular species seen on Western blots.

coprotein (Fig. 3*c*), and mass spectrometry demonstrated multiple molecular species consistent with variable glycosylation. Treatment of recombinant protein with peptide *N*-glycosidase F did not result in a perceptible decrease in molecular mass, although the intensity of glycoprotein staining was decreased (data not shown). Culture medium from tunicamycin-treated cells exhibited only a single punctin species as demonstrated by Western blotting (Fig. 3*d*). The difference (161 Da) between the LC-ESMS-observed masses of the major punctin molecular species (61,970 and 62,131 Da) is close to the in-chain chemical average mass of a oligosaccharide residue (Hex, 162). Minor molecular species were also apparent by LC-ESMS analysis, which differed by mass increments that approximated the in-chain chemical average mass of oligosaccharide residues (*e.g.* Hex, 162; HexNAc, 203; NeuAc, 291). For a further analysis, tryptic digests of the protein were examined by analytical LC-ESMS using stepped collision energy scanning to produce carbohydrate-specific marker ions. Glycopeptides were detected including molecular species with masses of 5881.4 ± 0.4 and 6171.2 ± 0.2 Da. The mass difference (289.8 Da) between these observed glycopeptides appears to correspond to the in-chain chemical average mass of *N*-acetylneuraminic acid (NeuAc, 291). Taken together, these data indicated that punctin is glycosylated, although specific glycopeptides have yet to be characterized fully. Approximately 65% of the amino acid sequence in punctin was identified by peptide mass mapping including the NH_2-terminal tryptic peptide (Glu^{29}-Arg^{47}), verifying that the target protein has been expressed. Based on the difference between the observed and calculated masses of intact punctin, the recombinant protein contains approximately 3–4% carbohydrate by weight.

During purification of punctin in the absence of protease inhibitors, additional components of ~40 and 20 kDa, respectively, were detected on Coomassie Blue-stained gels (data not shown). The 40-kDa band contained two molecular species with measured masses of 38,409 ± 115 and 39,456 ± 156 Da, respectively, as determined by MALDI-TOF MS. The NH_2-terminal sequencing of these bands yielded the same amino terminus as the full-length punctin. The ~20-kDa fragment exhibited an NH_2-terminal sequence ^{372}DLYHPL, indicating that the fragment is from the carboxyl terminus. The addition of 1 mM phenylmethylsulfonyl fluoride to culture medium effectively prevented this proteolysis, suggesting that it was effected by a serine protease.

Visualization of Punctin by Rotary Shadowing—Rotary shadowing of purified recombinant punctin demonstrated a hatchet-shaped or comma-shaped molecule 30–40 μm in length (Fig. 4). Punctin consists of a single globular domain of 10–20 μm in size with a short linear segment at one end. Most of the visualized protein was in monomeric form (Fig. 4). Occasional aggregates with the appearance of dimers and trimers were seen but have not yet been resolved in detail.

Expression and Localization of Punctin in Transfected COS-1 Cells—Transfected cells were stained without fixation or permeabilization and on ice (live staining) to prevent the detection of intracellular punctin or endocytosed antibody, respectively. Under these conditions, punctin was localized underneath the cells (*i.e.* adjacent to their ventral surface) in the substratum laid down on plastic. The staining pattern was punctate (Fig. 5, *a–d*) and was preferentially located toward the periphery of the cells (Fig. 5, *a*, *b*, and *d*) and under cellular processes (Fig. 5*c*). The punctin deposits were of submicron dimension, although fluorescent signals from closely located deposits were frequently merged suggesting larger aggregates. Transfected cells had minimal or no staining on the dorsal cell surface. Punctin was not seen in the substratum in areas not corresponding to the cells. If cells were detached with 10 mM EDTA prior to staining, "footprints" of transfected cells were retained on the substratum with a similar staining pattern as under intact

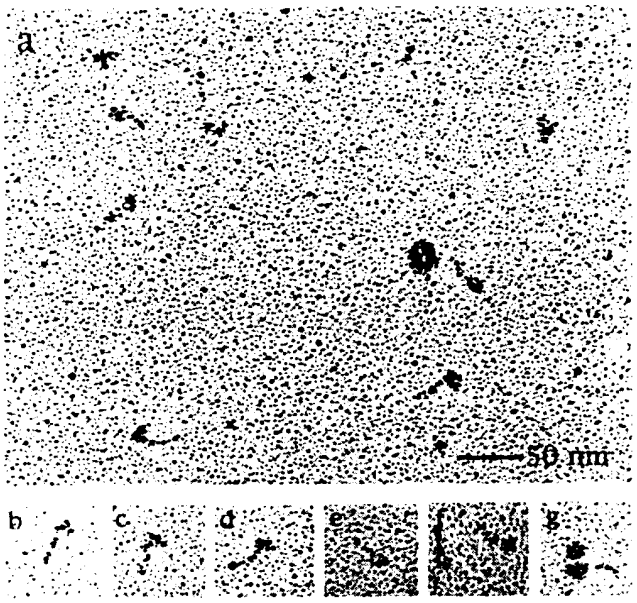

FIG. 4. **Rotary shadowing of recombinant punctin.** *a*, overview. *b–g*, images of individual punctin molecules. *Scale bar* in *panel a* indicates molecular dimensions in all panels.

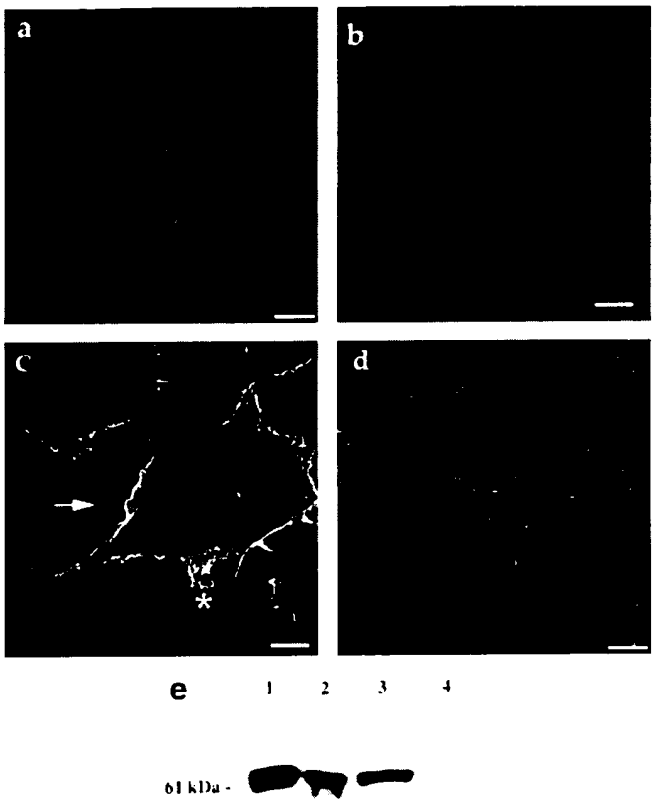

FIG. 5. *a–d*, confocal laser-scanning microscopy of COS-1 cells following transient transfection with *ADAMTSL1* expression constructs and immunocytochemistry. Untransfected cells are visible in *a* and *b*. Scale bar (10 μm) is shown at *lower right* of each panel. *a* and *b*, punctate staining of FLAG-tagged punctin (*red*) in nonpermeabilized cells visualized with anti-FLAG M2 antibody. Nuclei are *blue* 4′,6-diamidino-2-phenylindole. *c*, relationship of punctin staining (*red*) visualized with anti-FLAG M2 monoclonal antibody to cellular actin as visualized by phalloidin staining (*green*). The *asterisk* indicates a cellular protrusion that does not have underlying punctin, and the *arrow* indicates punctin immunolocalization without an overlying cellular process. *d*, relationship of punctin (*red*) visualized with anti-punctin antiserum 4113 to vinculin staining (*green*) as shown by confocal imaging and overlay of single-color images from a double-stained cell. *e*, Western blot analysis of cell lysates (*lane 1*), medium (*lane 2*), and ECM (*lane 3*) from transfected COS-1 cells using an anti-His tag monoclonal antibody. Cell lysates from untransfected COS-1 cells are shown in *lane 4*. Molecular mass is indicated on the *left*.

cells. Staining was seen in some areas not covered with cell processes. In other areas, there were cell processes without underlying punctin (Fig. 5*c*). We interpret this finding to result from cellular motility (*i.e.* withdrawal of existing processes and the formation of new ones). Identical results were obtained with anti-FLAG monoclonal antibody or antibody 4113. Fig 5, *a–c*, shows staining of FLAG-tagged protein using the FLAG M2 monoclonal antibody, and Fig. 5*d* shows staining with anti-punctin antiserum 4113. Similar staining patterns were seen whether cells were grown in the presence or absence of serum and using tagged or untagged proteins (data not shown).

Double staining for vinculin (Fig. 5*d*) or focal adhesion kinase (data not shown), components of focal contacts, indicated that punctin staining did not correspond to sites of focal contacts. No staining was visible in control experiments, *i.e.* in untransfected COS cells, cells transfected with vector alone, cells stained without a primary antibody, or cells stained with preimmune serum as control.

On Western blots, we found reactive protein bands of the expected size (58–60 kDa for untagged punctin and 62–64 kDa for the His-tagged or FLAG-tagged forms) in the medium, cell layer, and the underlying substratum or ECM of transfected COS-1 cells (Fig. 5*e*). In contrast, cells transfected with vector alone (Fig. 5*e*) or untransfected cells (data not shown) did not show a reactive band. As controls, preimmune serum from the rabbits in which anti-Punctin antibodies were generated did not produce immunoreactivity on Western blots (data not shown).

DISCUSSION

Punctin/ADAMTSL-1 Is a Novel ADAMTS-like Secreted Protein Belonging to a Distinct ADAMTSL Family of Proteins—In addition to missing the catalytic domain, the ADAMTS-like proteins (see below) do not possess disintegrin-like domains. This finding suggests that the disintegrin-like domain and catalytic domain may represent a functionally coupled protease domain in ADAMTS enzymes. Further evidence for this comes from the identification of other proteins with a predicted structure similar to punctin. Following the complete cloning of punctin/ADAMTSL-1, we became aware of a second such molecule encoded by the *KIAA0605* gene (GenBank™ accession number AB011177) that we designated as ADAMTSL-2 (23). We have cloned a third ADAMTS-like protein, ADAMTSL-3 (GenBank™ accession number AF237652).[3] Therefore, punctin belongs to a distinct protein family. ADAMTSL-2 and ADAMTSL-3 differ from punctin in their greater length (951 and 1690 amino acids, respectively) and also have more TS domains (6 and 10, respectively). These molecules will be described in greater detail in subsequent publications. In contrast to ADAMTSL-2 and ADAMTSL-3, which are quite widely expressed,[4] punctin/ADAMTSL-1 is selectively expressed in muscle.

Other secreted ECM molecules such as lacunin and papilin also contain the ancillary domains of the ADAMTS family in the precise order as punctin. However, punctin is more closely related to ADAMTSL-3 and some ADAMTS proteases than it is to mouse papilin (32% identity). Lacunin is a basement membrane glycoprotein in the moth *Manduca sexta* (24). Lacunin has the structure of ADAMTSL including seven TS modules as

[3] N. Moore, B. Anand-Apte, and S. Apte, unpublished data.
[4] S. Apte, unpublished data.

well as a single COOH-terminal protease and lacunin domain. In addition, it contains 13 repeats of a novel lagrin domain, 11 Kunitz inhibitor domains, 2 antistasin-like domains, 1 serine protease inhibitor domain, and 2 immunoglobulin domains. Lacunin localizes to the basal lamina of the moth wing (24). Papilin from *Drosophila melanogaster* may be an ortholog of *M. sexta* lacunin, because the two molecules are similar in their domain content, organization, and primary sequence. Papilin is also a basement membrane protein (25). Although these invertebrate proteins have numerous protease inhibitor domains, mammalian papilin contains substantially fewer such domains (25).

Characterization of Recombinant Punctin from Insect Cells—Our experimental data support the likelihood that recombinant punctin is disulfide-bonded. First, its electrophoretic mobility is greater under nonreducing conditions. Second, the punctin epitope is masked under nonreducing conditions. Third, rotary shadowing demonstrated a molecule with a specific and consistent conformation. Limited proteolysis within the linker peptide, connecting TS domains 2 and 3 assigned to the Tyr^{371}-Asp^{372} peptide bond (Fig. 1*b*) by a putative serine protease, indicates that there may be a proteolytically susceptible exposed region between the two disulfide-bonded TS domains. It is not yet known whether this is a physiologically relevant processing or whether it is an artifact that is unique to this expression system. The processing event releases the two COOH-terminal TS domains of punctin. Because proteolytically derived fragments of many secreted proteins have distinctive functions, it will be interesting to investigate whether specific functions are associated with the ~40- and ~20-kDa fragments.

A mass measurement of epitope-tagged recombinant punctin by MALDI-TOF MS and LC-ESMS revealed that purified punctin contained multiple species of higher than the predicted mass. Edman degradation indicated that all these species had the same amino terminus. Further MS analysis, glycoprotein staining, and culture in the presence of tunicamycin A confirm that punctin contains *N*-linked sugars but do not exclude the presence of *O*-linked sugar. Significant alteration of mobility was not seen after peptide *N*-glycosidase F treatment, suggesting that the *N*-linked carbohydrate may be resistant to complete enzymatic removal (26).

Rotary shadowing is useful for demonstrating the physical conformation of a molecule as well as the existence of oligomeric complexes (27–29). The data we have obtained for punctin are relevant to the ADAMTS, lacunin, and papilin. They can be extrapolated to represent the structure of the ancillary domains of an ADAMTS enzyme and the "papilin cassette" (25) and provide the first insight into the conformation of these domain assemblies. Many ECM proteins exist as oligomers. This observation may also be the case with punctin, because rotary shadowing electron microscopy and gel electrophoresis occasionally suggested the presence of dimers and trimers. We anticipate that rotary shadowing will be useful for future studies to investigate punctin oligomerization and interactions of punctin with putative ECM ligands.

Punctin Is an ECM Glycoprotein That Binds to the Cell Substratum in a Spatially Specific Manner—Nontransformed cells in culture require a substratum for attachment, spreading, and migration. The substratum present on an unmodified plastic tissue culture surface is derived from the cells themselves as well as from proteins in serum-supplemented culture medium (30–32). Quantitatively significant components of the cell substratum are laminin, fibronectin, vitronectin, collagen, tenascin, PG-M or versican (a chondroitin sulfate proteoglycan), perlecan (a heparan sulfate proteoglycan), hyaluronan, and tissue inhibitor of metalloproteases-3 (30–37). Punctin shares the subcellular distribution of molecules that do not generally co-localize with focal contacts (*e.g.* versican, hyaluronan, and tenascin) (31, 37). Because punctin is left behind in the ECM after cell detachment with EDTA, we conclude that when expressed in COS-1 cells, punctin binds a component of the ECM. Punctin in culture medium may reflect an excess of more than that which can bind to the substratum or indicate secretion from the free surface of the cell. Punctin does not bind to ECM between the cells, indicating that the punctin ligand is absent from these regions. Because similar staining was seen under serum-supplemented as well as under serum-free culture conditions, it is probable that the ECM binding partner of punctin is a molecule produced by COS-1 cells but not one derived from fetal bovine serum.

Significance of Punctin and the ADAMTS-like Family—Molecules comprising ancillary domains of metalloproteases may be generated in biological systems by proteolytic processing or through alternative splicing of protease genes. Brooks *et al.* (38) found that the proteolytically generated hemopexin domain of MMP-2 circulated in serum and bound to the integrin $\alpha_v\beta_3$. This MMP-2 fragment inhibited angiogenesis by preventing membrane targeting of MMP-2 (38). So far, there are no known examples of ADAMTS-like proteins generated as splice variants of *ADAMTS* genes. The discovery of punctin demonstrates for the first time the existence of molecules closely resembling the ancillary domains of ADAMTS that are generated as distinct gene products.

The resemblance of ADAMTSL to ADAMTS suggests a functional relationship between these two groups of molecules. From studies on ADAMTS-1 (39) and ADAMTS-2 (40), it is known that the ancillary domains are required to bind and cleave substrates. ADAMTSL may offer a potential mechanism of ADAMTS regulation via one of several possible mechanisms. As a result of noncompetitive inhibition of ADAMTS-2, an inhibitory role has been shown for *Drosophila* papilin (25). Another possibility is that punctin may compete with ADAMTS for its substrates and protect the substrates from cleavage. The isolated MMP-2 hemopexin domain represents one such example. In a second example, a truncated nonenzymatic version of ADAM-17 was shown to have a dominant negative effect on the activation of tumor necrosis factor-α (41). An intriguing possibility is that the ADAMTS-like proteins may be enhancers of the ADAMTS proteases. For example, the procollagen C-proteinase enhancer protein (42) contains two domains homologous to those found in the C-proteinase that are instrumental in binding to the carboxyl propeptide of procollagen I and enhancing its removal (43). Very little is currently known about the regulation of ADAMTS proteases following their activation, and it is possible that the ADAMTS-like proteins may provide a novel general principle of regulation.

Acknowledgments—We thank C. Kassuba for editing the manuscript, Vincent C. Hascall, Tom Tallant, Seng Hui Low, and Thomas Weimbs for helpful discussion, members of the Apte laboratory for critical reading of the manuscript, Satya Yadav for protein sequencing, Monique Ross for assistance with cloning, Karen West for amino acid analysis, and Judy Drazba for help with confocal imaging.

REFERENCES

1. Murphy, G., and Knauper, V. (1997) *Matrix Biol.* **15,** 511–518
2. Blobel, C. P. (1997) *Cell* **90,** 589–592
3. Kuno, K., Kanada, N., Nakashima, E., Fujiki, F., Ichimura, F., and Matsushima, K. (1997) *J. Biol. Chem.* **272,** 556–562
4. Hurskainen, T. L., Hirohata, S., Seldin, M. F., and Apte, S. S. (1999) *J. Biol. Chem.* **274,** 25555–25563
5. Wallon, U. M., and Overall, C. M. (1997) *J. Biol. Chem.* **272,** 7473–7481
6. Steffensen, B., Bigg, H. F., and Overall, C. M. (1998) *J. Biol. Chem.* **273,** 20622–20628
7. Bein, K., and Simons, M. (2000) *J. Biol. Chem.* **275,** 32167–32173
8. Bornstein, P., Armstrong, L. C., Hankenson, K. D., Kyriakides, T. R., and Yang, Z. (2000) *Matrix Biol.* **19,** 557–568

9. Overall, C. M., King, A. E., Sam, D. K., Ong, A. D., Lau, T. T., Wallon, U. M., De Clerck, Y. A., and Atherstone, J. (1999) *J. Biol. Chem.* **274,** 4421–4429
10. McQuibban, G. A., Gong, J. H., Tam, E. M., McCulloch, C. A., Clark-Lewis, I., and Overall, C. M. (2000) *Science* **289,** 1202–1206
11. Zhang, X. P., Kamata, T., Yokoyama, K., Puzon-McLaughlin, W., and Takada, Y. (1998) *J. Biol. Chem.* **273,** 7345–7350
12. Zolkiewska, A. (1999) *Exp. Cell Res.* **252,** 423–431
13. Kuno, K., and Matsushima, K. (1998) *J. Biol. Chem.* **273,** 13912–13917
14. Georgiadis, K. E., Hirohata, S., Seldin, M. F., and Apte, S. S. (1999) *Genomics* **62,** 312–315
15. Watson, M. L., and Seldin, M. F. (1994) *Methods Mol. Genet.* **5,** 369–387
16. Crabb, J. W., West, K. A., Dodson, W. S., and Hulmes, J. D. (1997) in *Current Protocols in Protein Science* (Coligan, J. E., Ploegh, H. L., Smith, J. A., and Speicher, D. W., eds) pp. 11.9.1–11.9.42, John Wiley & Sons, Inc., New York
17. Kapron, J. T., Hilliard, G. M., Lakins, J. N., Tenniswood, M. P., West, K. A., Carr, S. A., and Crabb, J. W. (1997) *Protein Sci.* **6,** 2120–2133
18. Crabb, J. W., Nie, Z., Chen, Y., Hulmes, J. D., West, K. A., Kapron, J. T., Ruuska, S. E., Noy, N., and Saari, J. C. (1998) *J. Biol. Chem.* **273,** 20712–20720
19. Sakai, L. Y., and Keene, D. R. (1994) *Methods Enzymol.* **245,** 29–52
20. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241
21. Gavel, Y., and von Heijne, G. (1990) *Protein Eng.* **3,** 433–442
22. Bornstein, P., and Sage, E. H. (1994) *Methods Enzymol.* **245,** 62–85
23. Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N., and Ohara, O. (1998) *DNA Res.* **5,** 31–39
24. Nardi, J. B., Martos, R., Walden, K. K., Lampe, D. J., and Robertson, H. M. (1999) *Insect Biochem. Mol. Biol.* **29,** 883–897
25. Kramerova, I. A., Kawaguchi, N., Fessler, L. I., Nelson, R. E., Chen, Y., Kramerov, A. A., Kusche-Gullberg, M., Kramer, J. M., Ackley, B. D., Sieron, A. L., Prockop, D. J., and Fessler, J. H. (2000) *Development* **127,** 5475–5485
26. Tretter, V., Altmann, F., and Marz, L. (1991) *Eur. J. Biochem.* **199,** 647–652
27. Holden, P., Meadows, R. S., Chapman, K. L., Grant, M. E., Kadler, K. E., and Briggs, M. D. (2001) *J. Biol. Chem.* **276,** 6046–6055
28. Maddox, B. K., Sakai, L. Y., Keene, D. R., and Glanville, R. W. (1989) *J. Biol. Chem.* **264,** 21381–21385
29. Ohn, A. I., Morgelin, M., Sasaki, T., Timpl, R., Heinegard, D., and Aspberg, A. (2001) *J. Biol. Chem.* **276,** 1253–1261
30. Hayman, E. G., Oldberg, A., Martin, G. R., and Ruoslahti, E. (1982) *J. Cell Biol.* **94,** 28–35
31. Hayman, E. G., Pierschbacher, M. D., Suzuki, S., and Ruoslahti, E. (1985) *Exp. Cell Res.* **160,** 245–258
32. Ohji, M., Sundar Raj, N., Hassell, J. R., and Thoft, R. A. (1994) *Invest. Ophthalmol. Visual Sci.* **35,** 479–485
33. Snow, A. D., Sekiguchi, R., Nochlin, D., Fraser, P., Kimata, K., Mizutani, A., Arai, M., Schreier, W. A., and Morgan, D. G. (1994) *Neuron* **12,** 219–234
34. Singer, I. I., II, Scott, S., Kawka, D. W., and Hassell, J. R. (1987) *Exp. Cell Res.* **173,** 558–571
35. Isogai, Z., Shinomura, T., Yamakawa, N., Takeuchi, J., Tsuji, T., Heinegard, D., and Kimata, K. (1996) *Cancer Res.* **56,** 3902–3908
36. Staskus, P. W., Masiarz, F. R., Pallanck, L. J., and Hawkes, S. P. (1991) *J. Biol. Chem.* **266,** 449–454
37. Yamagata, M., Saga, S., Kato, M., Bernfield, M., and Kimata, K. (1993) *J. Cell Sci.* **106,** 55–65
38. Brooks, P. C., Silletti, S., von Schalscha, T. L., Friedlander, M., and Cheresh, D. A. (1998) *Cell* **92,** 391–400
39. Kuno, K., Terashima, Y., and Matsushima, K. (1999) *J. Biol. Chem.* **274,** 18821–18826
40. Cohge, A., Sieron, A. L., Li, S. W., Schwarze, U., Petty, E., Wertelecki, W., Wilcox, W., Krakow, D., Cohn, D. H., Reardon, W., Byers, P. H., Lapiere, C. M., Prockop, D. J., and Nusgens, B. V. (1999) *Am. J. Hum. Genet.* **65,** 308–317
41. Solomon, K. A., Pesti, N., Wu, G., and Newton, R. C. (1999) *J. Immunol.* **163,** 4105–4108
42. Takahara, K., Kessler, E., Biniaminov, L., Brusel, M., Eddy, R. L., Jani-Sait, S., Shows, T. B., and Greenspan, D. S. (1994) *J. Biol. Chem.* **269,** 26280–26285
43. Hulmes, D. J., Mould, A. P., and Kessler, E. (1997) *Matrix Biol.* **16,** 41–45

```
>AC027807.6.1.181327
          Length = 181327

 Score = 2117 bits (1068), Expect = 0.0
 Identities = 1071/1072 (99%)
 Strand = Plus / Minus


Query: 2645   aaatcaaatcagagatgaagacaaaacttggtgagcagggtccgcagatcctcagtgtcc 2704
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104240 aaatcaaatcagagatgaagacaaaacttggtgagcagggtccgcagatcctcagtgtcc 104181


Query: 2705   agagagtctacattcagacaagggaagagaagcgtattaacctgaccattggtagcagag 2764
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104180 agagagtctacattcagacaagggaagagaagcgtattaacctgaccattggtagcagag 104121


Query: 2765   cctatttgctgcccaacacatccgtgattattaagtgccccgtgcgacgattccagaaat 2824
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104120 cctatttgctgcccaacacatccgtgattattaagtgccccgtgcgacgattccagaaat 104061


Query: 2825   ctctgatccagtgggagaaggatggccgttgcctgcagaactccaaacggcttggcatca 2884
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 104060 ctctgatccagtgggagaaggatggccgttgcctgcagaactccaaacggcttggcatca 104001


Query: 2885   ccaagtcaggctcactaaaatccacggtcttgctgcccccgacatcggcgtgtaccggt 2944
               |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Sbjct: 104000 ccaagtcaggctcactaaaatccatggtcttgctgcccccgacatcggcgtgtaccggt 103941


Query: 2945   gcattgcaggctctgcacaggaaacagttgtgctcaagctcattggtactgacaaccggc 3004
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103940 gcattgcaggctctgcacaggaaacagttgtgctcaagctcattggtactgacaaccggc 103881


Query: 3005   tcatcgcacgcccagccctcagggagcctatgagggaatatcctgggatggaccacagcg 3064
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103880 tcatcgcacgcccagccctcagggagcctatgagggaatatcctgggatggaccacagcg 103821


Query: 3065   aagccaatagtttgggagtcacatggcacaaaatgaggcaaatgtggaataacaaaaatg 3124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103820 aagccaatagtttgggagtcacatggcacaaaatgaggcaaatgtggaataacaaaaatg 103761


Query: 3125   acctttatctggatgatgaccacattagtaaccagcctttcttgagagctctgttaggcc 3184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103760 acctttatctggatgatgaccacattagtaaccagcctttcttgagagctctgttaggcc 103701


Query: 3185   actgcagcaattctgcaggaagcaccaactcctgggagttgaagaataagcagtttgaag 3244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103700 actgcagcaattctgcaggaagcaccaactcctgggagttgaagaataagcagtttgaag 103641
```

```
Query: 3245   cagcagttaaacaaggagcatatagcatggatacagcccagtttgatgagctgataagaa 3304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103640 cagcagttaaacaaggagcatatagcatggatacagcccagtttgatgagctgataagaa 103581


Query: 3305   acatgagtcagctcatggaaaccggagaggtcagcgatgatcttgcgtcccagctgatat 3364
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103580 acatgagtcagctcatggaaaccggagaggtcagcgatgatcttgcgtcccagctgatat 103521


Query: 3365   atcagctggtggccgaattagccaaggcacagccaacacacatgcagtggcggggcatcc 3424
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103520 atcagctggtggccgaattagccaaggcacagccaacacacatgcagtggcggggcatcc 103461


Query: 3425   aggaagagacacctcctgctgctcagctcagaggggaaacagggagtgtgtcccaaagct 3484
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103460 aggaagagacacctcctgctgctcagctcagaggggaaacagggagtgtgtcccaaagct 103401


Query: 3485   cgcatgcaaaaaactcaggcaagctgacattcaagccgaaaggacctgttctcatgaggc 3544
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103400 cgcatgcaaaaaactcaggcaagctgacattcaagccgaaaggacctgttctcatgaggc 103341


Query: 3545   aaagccaacctccctcaatttcatttaataaaacaataaattccaggattggaaatacag 3604
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103340 aaagccaacctccctcaatttcatttaataaaacaataaattccaggattggaaatacag 103281


Query: 3605   tatacattacaaaaaggacagaggtcatcaatatactgtgtgaccttattacccccagtg 3664
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103280 tatacattacaaaaaggacagaggtcatcaatatactgtgtgaccttattacccccagtg 103221


Query: 3665   aggccacatatacatggaccaaggatggaaccttgttacagccctcagtaaa 3716
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 103220 aggccacatatacatggaccaaggatggaaccttgttacagccctcagtaaa 103169



 Score =  547 bits (276), Expect = e-152
 Identities = 285/288 (98%)
 Strand = Plus / Minus


Query: 1701   gttcattccagaaccctggtcagcctgcagtaccacgtgtgggccaggtgtgcaggtccg 1760
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 173421 gttcattccagaaccctggtcagcctgcagtaccacgtgtgggccgggtgtgcaggtccg 173362


Query: 1761   cgaggtgaagtgccgtgtgctcctcacattcacgcagactgagactgagctgcccgagga 1820
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 173361 tgaggtgaagtgccgtgtgctcctcacattcacgcagactgagactgagctgcccgagga 173302


Query: 1821   agagtgtgaaggccccaagctgcccaccgaacggccctgcctcctggaagcatgtgatga 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 173301 agagtgtgaaggccccaagctgcccaccgaacggccctgcctcctggaagcatgtgatga 173242
```

```
Query: 1881   gagcccggcctcccgagagctagacatccctctccctgaggacagtgagacgacttacga 1940

Sbjct: 173241 gagcccggcctcccgagagctagacatccctctccctgaggacagtgagacgacttacga 173182


Query: 1941   ctgggagtacgctgggttcaccccttgcacagcaacatgcttgggagg 1988

Sbjct: 173181 ctgggagtacgctgggttcaccccttgcacagcaacatgcgtgggagg 173134


 Score =  432 bits (218), Expect = e-118
 Identities = 218/218 (100%)
 Strand = Plus / Minus


Query: 4753  agacccacccttaagaaggaactgcacatcaggggcctgtgatgtgtgttggcacacaggc 4812

Sbjct: 49742 agacccacccttaagaaggaactgcacatcaggggcctgtgatgtgtgttggcacacaggc 49683


Query: 4813  cctttggaagccctgtacagcagcctgtggcaggggtttccagtctcggaaagtcgactgt 4872

Sbjct: 49682 cctttggaagccctgtacagcagcctgtggcaggggtttccagtctcggaaagtcgactgt 49623


Query: 4873  atccacacaaggagttgcaaacctgtggccaagagacactgtgtacagaaaaagaaacca 4932

Sbjct: 49622 atccacacaaggagttgcaaacctgtggccaagagacactgtgtacagaaaaagaaacca 49563


Query: 4933  atttcctggcggcactgtcttgggccctcctgtgatag 4970

Sbjct: 49562 atttcctggcggcactgtcttgggccctcctgtgatag 49525


 Score =  398 bits (201), Expect = e-107
 Identities = 201/201 (100%)
 Strand = Plus / Minus


Query: 4456  aggtggttcacaagtgtgtggtcacagtgctctgtgtcttgcggtgaaggataccacagt 4515

Sbjct: 61277 aggtggttcacaagtgtgtggtcacagtgctctgtgtcttgcggtgaaggataccacagt 61218


Query: 4516  cggcaggtgacgtgcaagcggacaaaagccaatggaactgtgcaggtggtgtctccaaga 4575

Sbjct: 61217 cggcaggtgacgtgcaagcggacaaaagccaatggaactgtgcaggtggtgtctccaaga 61158


Query: 4576  gcatgtgcccctaaagaccggcctctgggaagaaaaccatgttttggtcatccatgtgtt 4635

Sbjct: 61157 gcatgtgcccctaaagaccggcctctgggaagaaaaccatgttttggtcatccatgtgtt 61098
```

```
Query: 4636  cagtgggaaccagggaaccgg 4656
              ||||||||||||||||||||
Sbjct: 61097 cagtgggaaccagggaaccgg 61077
```

Score = 387 bits (195), Expect = e-104
Identities = 195/195 (100%)
Strand = Plus / Minus

```
Query: 2116   aggtggcatgtgggctcttgggggccctgctcagctacctgtggagttggaattcagacc 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 143919 aggtggcatgtgggctcttgggggccctgctcagctacctgtggagttggaattcagacc 143860

Query: 2176   cgagatgtgtactgcctgcacccaggggagacccctgcccctcctgaggagtgccgagat 2235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 143859 cgagatgtgtactgcctgcacccaggggagacccctgcccctcctgaggagtgccgagat 143800

Query: 2236   gaaaagccccatgctttacaagcatgcaatcagtttgactgccctcctggctggcacatt 2295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 143799 gaaaagccccatgctttacaagcatgcaatcagtttgactgccctcctggctggcacatt 143740

Query: 2296   gaagaatggcagcag 2310
              ||||||||||||||
Sbjct: 143739 gaagaatggcagcag 143725
```

Score = 373 bits (188), Expect = e-100
Identities = 188/188 (100%)
Strand = Plus / Minus

```
Query: 4272  agagccttttgggagcctggtaactggtcacattgttctgccacctgtggtcatttggg 4331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 65105 agagccttttgggagcctggtaactggtcacattgttctgccacctgtggtcatttggg 65046

Query: 4332  agcccgcattcagagaccccagtgtgtgatggccaatgggcaggaagtgagtgaggccct 4391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 65045 agcccgcattcagagaccccagtgtgtgatggccaatgggcaggaagtgagtgaggccct 64986

Query: 4392  gtgtgatcacctccagaagccactggctgggtttgagccctgtaacatccgggactgccc 4451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 64985 gtgtgatcacctccagaagccactggctgggtttgagccctgtaacatccgggactgccc 64926

Query: 4452  agcgaggt 4459
              ||||||||
Sbjct: 64925 agcgaggt 64918
```

```
 Score =  371 bits (187), Expect = 3e-99
 Identities = 187/187 (100%)
 Strand = Plus / Minus


Query: 3970  aaaggtgtcctcagcctaatataacttggttgaagagaggaggatctctgagtggcaat 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71975 aaaggtgtcctcagcctaatataacttggttgaagagaggaggatctctgagtggcaat 71916


Query: 4030  gtttccttgcttttcaatggatccctgttgttgcagaatgtttcccttgaaaatgaagga 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71915 gtttccttgcttttcaatggatccctgttgttgcagaatgtttcccttgaaaatgaagga 71856


Query: 4090  acctacgtctgcatagccaccaatgctcttggaaaggcagtggcaacatctgtactccac 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 71855 acctacgtctgcatagccaccaatgctcttggaaaggcagtggcaacatctgtactccac 71796


Query: 4150  ttgctgg 4156
              |||||||
Sbjct: 71795 ttgctgg 71789


 Score =  361 bits (182), Expect = 3e-96
 Identities = 182/182 (100%)
 Strand = Plus / Minus


Query: 2309   agtgttccaggacttgtggcgggggaactcagaacagaagagtcacctgtcggcagctgc 2368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 143612 agtgttccaggacttgtggcgggggaactcagaacagaagagtcacctgtcggcagctgc 143553


Query: 2369   taacggatggcagctttttgaatctctcagatgaattgtgccaaggacccaaggcatcgt 2428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 143552 taacggatggcagctttttgaatctctcagatgaattgtgccaaggacccaaggcatcgt 143493


Query: 2429   ctcacaagtcctgtgccaggacagactgtcctccacatttagctgtgggagactggtcga 2488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 143492 ctcacaagtcctgtgccaggacagactgtcctccacatttagctgtgggagactggtcga 143433


Query: 2489   ag 2490
              ||
Sbjct: 143432 ag 143431
```

```
 Score =  301 bits (152), Expect = 2e-78
 Identities = 155/156 (99%)
 Strand = Plus   Minus


Query: 2489   agtgttctgtcagttgtggtgttggaatccagagaagaaagcaggtgtgtcaaaggctgg 2548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 116031 agtgttctgtcagttgtggtgttggaatccagagaagaaagcaggtgtgtcaaaggctgg 115972


Query: 2549   cagccaaaggtcggcgcatccccctcagtgagatgatgtgcagggatctaccagggttcc 2608
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct: 115971 cagccaaaggtcggcgcatccccctcagtgagatgatgtgcagggatctaccagggctcc 115912


Query: 2609   ctcttgtaagatcttgccagatgcctgagtgcagta 2644
              ||||||||||||||||||||||||||||||||||||
Sbjct: 115911 ctcttgtaagatcttgccagatgcctgagtgcagta 115876



 Score =  266 bits (134), Expect = 1e-67
 Identities = 134/134 (100%)
 Strand = Plus / Minus


Query: 1986   aggccatcaagaagccatagcagtgtgcttacatatccagacccagcagacagtcaatga 2045
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 162611 aggccatcaagaagccatagcagtgtgcttacatatccagacccagcagacagtcaatga 162552


Query: 2046   cagcttgtgtgatatggtccaccgtcctccagccatgagccaggcctgtaacacagagcc 2105
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 162551 cagcttgtgtgatatggtccaccgtcctccagccatgagccaggcctgtaacacagagcc 162492


Query: 2106   ctgtccccccaggt 2119
              ||| ||||||||||
Sbjct: 162491 ctgtccccccaggt 162478



 Score =  264 bits (133), Expect = 5e-67
 Identities = 133/133 (100%)
 Strand = Plus / Minus


Query: 3843  agaggcacctgtcatcttgtctgttgaaagaaatatcaccaaaccagagcacaaccatct 3902
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 95429 agaggcacctgtcatcttgtctgttgaaagaaatatcaccaaaccagagcacaaccatct 95370


Query: 3903  gtctgttgtggttggaggcatcgtggaggcagcccttggagcaaacgtgacaatccgatg 3962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 95369 gtctgttgtggttggaggcatcgtggaggcagcccttggagcaaacgtgacaatccgatg 95310
```

```
Query: 3963  tcctgtaaaaggt 3975
              |||||||||||
Sbjct: 95309 tcctgtaaaaggt 95297
```

```
 Score =  254 bits (128), Expect = 5e-64
 Identities = 128/128 (100%)
 Strand = Plus / Minus
```

```
Query: 3717  aataattttggatggaactgggaagatacagatacagaatcctacaaggaaagaacaagg 3776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 97812 aataattttggatggaactgggaagatacagatacagaatcctacaaggaaagaacaagg 97763
```

```
Query: 3777  catatatgaatgttctgtagctaatcatcttggttcagatgtggaaagttcttctgtgct 3836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 97762 catatatgaatgttctgtagctaatcatcttggttcagatgtggaaagttcttctgtgct 97703
```

```
Query: 3837  gtatgcag 3844
              ||||||||
Sbjct: 97702 gtatgcag 97695
```

```
 Score =  234 bits (118), Expect = 4e-58
 Identities = 118/118 (100%)
 Strand = Plus / Minus
```

```
Query: 4156  gaacgaagatggccagagagtagaatcgtatttctgcaaggacataaaaagtacattctc 4215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 70131 gaacgaagatggccagagagtagaatcgtatttctgcaaggacataaaaagtacattctc 70072
```

```
Query: 4216  caggcaaccaacactagaaccaacagcaatgacccaacaggagaaccccccgcctcaag 4273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 70071 caggcaaccaacactagaaccaacagcaatgacccaacaggagaaccccccgcctcaag 70014
```

```
 Score =  212 bits (107), Expect = 2e-51
 Identities = 107/107 (100%)
 Strand = Plus / Minus
```

```
Query: 4970  gagactgcacagacacaactcactactgtatgtttgtaaaacatcttaatttgtgttctc 5029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 48813 gagactgcacagacacaactcactactgtatgtttgtaaaacatcttaatttgtgttctc 48754
```

```
Query: 5030  tagaccgctacaaacaaaggtgctgccagtcatgtcaagagggataa 5076
              |||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 48755 tagaccgctacaaacaaaggtgctgccagtcatgtcaagagggataa 48707
```

```
 Score =  196 bits (99), Expect = 1e-46
 Identities = 99/99 (100%)
 Strand = Plus / Minus


Query: 4656  gtgtcctggacgttgcatgggccgtgctgtgaggatgcagcagcgtcacacagcttgtca 4715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 55179 gtgtcctggacgttgcatgggccgtgctgtgaggatgcagcagcgtcacacagcttgtca 55120


Query: 4716  acacaacagctctgactccaactgtgatgacagaaagag 4754
              |||||||||||||||||||||||||||||||||||||
Sbjct: 55119 acacaacagctctgactccaactgtgatgacagaaagag 55081



>AC012684.5.1.164879
          Length = 164879

 Score = 2117 bits (1068), Expect = 0.0
 Identities = 1071/1072 (99%)
 Strand = Plus / Minus


Query: 2645  aaatcaaatcagagatgaagacaaaacttggtgagcagggtccgcagatcctcagtgtcc 2704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 33829 aaatcaaatcagagatgaagacaaaacttggtgagcagggtccgcagatcctcagtgtcc 33770


Query: 2705  agagagtctacattcagacaagggaagagaagcgtattaacctgaccattggtagcagag 2764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 33769 agagagtctacattcagacaagggaagagaagcgtattaacctgaccattggtagcagag 33710


Query: 2765  cctatttgctgcccaacacatccgtgattattaagtgccccgtgcgacgattccagaaat 2824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 33709 cctatttgctgcccaacacatccgtgattattaagtgccccgtgcgacgattccagaaat 33650


Query: 2825  ctctgatccagtgggagaaggatggccgttgcctgcagaactccaaacggcttggcatca 2884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 33649 ctctgatccagtgggagaaggatggccgttgcctgcagaactccaaacggcttggcatca 33590


Query: 2885  ccaagtcaggctcactaaaaatccacggtcttgctgcccccgacatcggcgtgtaccggt 2944
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct: 33589 ccaagtcaggctcactaaaaatccatggtcttgctgcccccgacatcggcgtgtaccggt 33530


Query: 2945  gcattgcaggctctgcacaggaaacagttgtgctcaagctcattggtactgacaaccggc 3004
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 33529 gcattgcaggctctgcacaggaaacagttgtgctcaagctcattggtactgacaaccggc 33470


Query: 3005  tcatcgcacgcccagccctcagggagcctatgagggaatatcctgggatggaccacagcg 3064
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 33469 tcatcgcacgcccagccctcagggagcctatgagggaatatcctgggatggaccacagcg 33410
```


Query: 3065 aagccaatagtttgggagtcacatggcacaaaatgaggcaaatgtggaataacaaaaatg 3124
Sbjct: 33409 aagccaatagtttgggagtcacatggcacaaaatgaggcaaatgtggaataacaaaaatg 33350
Query: 3125 acctttatctggatgatgaccacattagtaaccagcctttcttgagagctctgttaggcc 3184
Sbjct: 33349 acctttatctggatgatgaccacattagtaaccagcctttcttgagagctctgttaggcc 33290
Query: 3185 actgcagcaattctgcaggaagcaccaactcctgggagttgaagaataagcagtttgaag 3244
Sbjct: 33289 actgcagcaattctgcaggaagcaccaactcctgggagttgaagaataagcagtttgaag 33230
Query: 3245 cagcagttaaacaaggagcatatagcatggatacagcccagtttgatgagctgataagaa 3304
Sbjct: 33229 cagcagttaaacaaggagcatatagcatggatacagcccagtttgatgagctgataagaa 33170
Query: 3305 acatgagtcagctcatggaaaccggagaggtcagcgatgatcttgcgtcccagctgatat 3364
Sbjct: 33169 acatgagtcagctcatggaaaccggagaggtcagcgatgatcttgcgtcccagctgatat 33110
Query: 3365 atcagctggtggccgaattagccaaggcacagccaacacacatgcagtggcggggcatcc 3424
Sbjct: 33109 atcagctggtggccgaattagccaaggcacagccaacacacatgcagtggcggggcatcc 33050
Query: 3425 aggaagagacacctcctgctgctcagctcagaggggaaacagggagtgtgtcccaaagct 3484
Sbjct: 33049 aggaagagacacctcctgctgctcagctcagaggggaaacagggagtgtgtcccaaagct 32990
Query: 3485 cgcatgcaaaaaactcaggcaagctgacattcaagccgaaaggacctgttctcatgaggc 3544
Sbjct: 32989 cgcatgcaaaaaactcaggcaagctgacattcaagccgaaaggacctgttctcatgaggc 32930
Query: 3545 aaagccaacctccctcaatttcatttaataaaacaataaattccaggattggaaatacag 3604
Sbjct: 32929 aaagccaacctccctcaatttcatttaataaaacaataaattccaggattggaaatacag 32870
Query: 3605 tatacattacaaaaaggacagaggtcatcaatatactgtgtgaccttattacccccagtg 3664
Sbjct: 32869 tatacattacaaaaaggacagaggtcatcaatatactgtgtgaccttattacccccagtg 32810
Query: 3665 aggccacatatacatggaccaaggatggaaccttgttacagccctcagtaaa 3716
Sbjct: 32809 aggccacatatacatggaccaaggatggaaccttgttacagccctcagtaaa 32758

```
 Score =  547 bits (276), Expect = e-152
 Identities = 285/288 (98%)
 Strand = Plus / Minus


Query: 1701   gttcattccagaaccctggtcagcctgcagtaccacgtgtgggccaggtgtgcaggtccg 1760
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct: 103010 gttcattccagaaccctggtcagcctgcagtaccacgtgtgggccgggtgtgcaggtccg 102951


Query: 1761   cgaggtgaagtgccgtgtgctcctcacattcacgcagactgagactgagctgcccgagga 1820
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct: 102950 tgaggtgaagtgccgtgtgctcctcacattcacgcagactgagactgagctgcccgagga 102891


Query: 1821   agagtgtgaaggccccaagctgcccaccgaacggccctgcctcctggaagcatgtgatga 1880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 102890 agagtgtgaaggccccaagctgcccaccgaacggccctgcctcctggaagcatgtgatga 102831


Query: 1881   gagcccggcctcccgagagctagacatcctctccctgaggacagtgagacgacttacga 1940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 102830 gagcccggcctcccgagagctagacatcctctccctgaggacagtgagacgacttacga 102771


Query: 1941   ctgggagtacgctgggttcacccccttgcacagcaacatgcttgggagg 1988
              ||||||||||||||||||||||||||||||||||||| |||||||
Sbjct: 102770 ctgggagtacgctgggttcacccccttgcacagcaacatgcgtgggagg 102723


 Score =  408 bits (206), Expect = e-110
 Identities = 206/206 (100%)
 Strand = Plus / Minus


Query: 1262   gctgggaacataatccttggactgcatgttccgtgtcctgtggaggagggattcagagac 1321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 123415 gctgggaacataatccttggactgcatgttccgtgtcctgtggaggagggattcagagac 123356


Query: 1322   ggagctttgtgtgtgtagaggaatccatgcatggagagatattgcaggtggaagaatgga 1381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 123355 ggagctttgtgtgtgtagaggaatccatgcatggagagatattgcaggtggaagaatgga 123296


Query: 1382   agtgcatgtacgcacccaaacccaaggttatgcaaacttgtaatctgtttgattgcccca 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 123295 agtgcatgtacgcacccaaacccaaggttatgcaaacttgtaatctgtttgattgcccca 123236


Query: 1442   agtggattgccatggagtggtctcag 1467
              |||||||||||||||||||||||||
Sbjct: 123235 agtggattgccatggagtggtctcag 123210
```

```
 Score =  387 bits (195), Expect = e-104
 Identities = 195/195 (100%)
 Strand = Plus / Minus


Query: 2116  aggtggcatgtgggctcttgggggccctgctcagctacctgtggagttggaattcagacc 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 73508 aggtggcatgtgggctcttgggggccctgctcagctacctgtggagttggaattcagacc 73449


Query: 2176  cgagatgtgtactgcctgcacccaggggagaccccctgccccctcctgaggagtgccgagat 2235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 73448 cgagatgtgtactgcctgcacccaggggagaccccctgccccctcctgaggagtgccgagat 73389


Query: 2236  gaaaagccccatgctttacaagcatgcaatcagtttgactgccctcctggctggcacatt 2295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 73388 gaaaagccccatgctttacaagcatgcaatcagtttgactgccctcctggctggcacatt 73329


Query: 2296  gaagaatggcagcag 2310
              ||||||||||||||
Sbjct: 73328 gaagaatggcagcag 73314


 Score =  371 bits (187), Expect = 3e-99
 Identities = 187/187 (100%)
 Strand = Plus / Minus


Query: 3970 aaaggtgtccctcagcctaatataacttggttgaagagaggaggatctctgagtggcaat 4029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1564 aaaggtgtccctcagcctaatataacttggttgaagagaggaggatctctgagtggcaat 1505


Query: 4030 gtttccttgcttttcaatggatccctgttgttgcagaatgtttcccttgaaaatgaagga 4089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1504 gtttccttgcttttcaatggatccctgttgttgcagaatgtttcccttgaaaatgaagga 1445


Query: 4090 acctacgtctgcatagccaccaatgctcttggaaaggcagtggcaacatctgtactccac 4149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 1444 acctacgtctgcatagccaccaatgctcttggaaaggcagtggcaacatctgtactccac 1385


Query: 4150 ttgctgg 4156
          |||||||
Sbjct: 1384 ttgctgg 1378
```

```
 Score =  361 bits (182), Expect = 3e-96
 Identities = 182/182 (100%)
 Strand = Plus / Minus


Query: 2309  agtgttccaggacttgtggcgggggaactcagaacagaagagtcacctgtcggcagctgc 2368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 73201 agtgttccaggacttgtggcgggggaactcagaacagaagagtcacctgtcggcagctgc 73142


Query: 2369  taacggatggcagctttttgaatctctcagatgaattgtgccaaggacccaaggcatcgt 2428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 73141 taacggatggcagctttttgaatctctcagatgaattgtgccaaggacccaaggcatcgt 73082


Query: 2429  ctcacaagtcctgtgccaggacagactgtcctccacatttagctgtgggagactggtcga 2488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 73081 ctcacaagtcctgtgccaggacagactgtcctccacatttagctgtgggagactggtcga 73022


Query: 2489  ag 2490
              ||
Sbjct: 73021 ag 73020


 Score =  305 bits (154), Expect = 1e-79
 Identities = 157/158 (99%)
 Strand = Plus / Minus


Query: 803    ttattgaatcaaaaacacttcaaggaagcaaaggagaacacagctttaacagccccggcg 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 145296 ttattgaatcaaaaacacttcaaggaagcaaaggagaacacagctttaacagccccggcg 145237


Query: 863    tctttgtcgtagaaaacacaacagtggaatttcagaggggctccgagaggcaaacttttta 922
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 145236 tctttctcgtagaaaacacaacagtggaatttcagaggggctccgagaggcaaacttttta 145177


Query: 923    agattccaggacctctgatggctgatttcatcttcaag 960
              |||||||||||||||||||||||||||||||||||||
Sbjct: 145176 agattccaggacctctgatggctgatttcatcttcaag 145139


 Score =  301 bits (152), Expect = 2e-78
 Identities = 155/156 (99%)
 Strand = Plus / Minus


Query: 2489  agtgttctgtcagttgtggtgttggaatccagagaagaaagcaggtgtgtcaaaggctgg 2548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 45620 agtgttctgtcagttgtggtgttggaatccagagaagaaagcaggtgtgtcaaaggctgg 45561
```

```
Query: 2549  cagccaaaggtcggcgcatccccctcagtgagatgatgtgcagggatctaccagggttcc 2608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct: 45560 cagccaaaggtcggcgcatccccctcagtgagatgatgtgcagggatctaccagggctcc 45501

Query: 2609  ctcttgtaagatcttgccagatgcctgagtgcagta 2644
              ||||||||||||||||||||||||||||||||||||
Sbjct: 45500 ctcttgtaagatcttgccagatgcctgagtgcagta 45465
```

Score = 297 bits (150), Expect = 4e-77
Identities = 150/150 (100%)
Strand = Plus / Minus

```
Query: 1466   agtgcacagtgacttgtggccgagggttacggtaccgggttgttctgtgtattaaccacc 1525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 118242 agtgcacagtgacttgtggccgagggttacggtaccgggttgttctgtgtattaaccacc 118183

Query: 1526   gcggagagcatgttggggctgcaatccacaactgaagttacacatcaaagaagaatgtg 1585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 118182 gcggagagcatgttggggctgcaatccacaactgaagttacacatcaaagaagaatgtg 118123

Query: 1586   tcattcccatcccgtgttataaaccaaaag 1615
              ||||||||||||||||||||||||||||||
Sbjct: 118122 tcattcccatcccgtgttataaaccaaaag 118093
```

Score = 280 bits (141), Expect = 8e-72
Identities = 141/141 (100%)
Strand = Plus / Minus

```
Query: 1071   aggttatcagctcaattctgctgaatgtgtggatatccgcttgaagagggtagttcctga 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 125991 aggttatcagctcaattctgctgaatgtgtggatatccgcttgaagagggtagttcctga 125932

Query: 1131   ccattattgtcactactaccctgaaaatgtaaaaccaaaaccaaaactgaaggaatgcag 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 125931 ccattattgtcactactaccctgaaaatgtaaaaccaaaaccaaaactgaaggaatgcag 125872

Query: 1191   catggatccctgcccatcaag 1211
              |||||||||||||||||||||
Sbjct: 125871 catggatccctgcccatcaag 125851
```

```
 Score =  266 bits (134), Expect = 1e-67
 Identities = 134/134 (100%)
 Strand = Plus / Minus


Query: 1986  aggccatcaagaagccatagcagtgtgcttacatatccagacccagcagacagtcaatga 2045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 92200 aggccatcaagaagccatagcagtgtgcttacatatccagacccagcagacagtcaatga 92141


Query: 2046  cagcttgtgtgatatggtccaccgtcctccagccatgagccaggcctgtaacacagagcc 2105
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 92140 cagcttgtgtgatatggtccaccgtcctccagccatgagccaggcctgtaacacagagcc 92081


Query: 2106  ctgtcccccccaggt 2119
              |||||||||||||||
Sbjct: 92080 ctgtcccccccaggt 92067


 Score =  264 bits (133), Expect = 5e-67
 Identities = 133/133 (100%)
 Strand = Plus / Minus


Query: 3843  agaggcacctgtcatcttgtctgttgaaagaaatatcaccaaaccagagcacaaccatct 3902
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25018 agaggcacctgtcatcttgtctgttgaaagaaatatcaccaaaccagagcacaaccatct 24959


Query: 3903  gtctgttgtggttggaggcatcgtggaggcagcccttggagcaaacgtgacaatccgatg 3962
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 24958 gtctgttgtggttggaggcatcgtggaggcagcccttggagcaaacgtgacaatccgatg 24899


Query: 3963  tcctgtaaaaggt 3975
              |||||||||||||
Sbjct: 24898 tcctgtaaaaggt 24886


 Score =  254 bits (128), Expect = 5e-64
 Identities = 128/128 (100%)
 Strand = Plus / Minus


Query: 3717  aataattttggatggaactgggaagatacagatacagaatcctacaaggaaagaacaagg 3776
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27411 aataattttggatggaactgggaagatacagatacagaatcctacaaggaaagaacaagg 27352


Query: 3777  catatatgaatgttctgtagctaatcatcttggttcagatgtggaaagttcttctgtgct 3836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27351 catatatgaatgttctgtagctaatcatcttggttcagatgtggaaagttcttctgtgct 27292
```

```
Query: 3837  gtatgcag 3844
              ||||||||
Sbjct: 27291 gtatgcag 27284
```

```
 Score =  230 bits (116), Expect = 7e-57
 Identities = 116/116 (100%)
 Strand = Plus / Minus

Query: 959    agaccaggtacactgcagccaaagacagcgtggttcagttcttcttttaccagcccatca 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 130999 agaccaggtacactgcagccaaagacagcgtggttcagttcttcttttaccagcccatca 130940

Query: 1019   gtcatcagtggagacaaactgacttctttcctgcactgtgacgtgtggaggaggt 1074
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 130939 gtcatcagtggagacaaactgacttctttcctgcactgtgacgtgtggaggaggt 130884
```

```
 Score =  176 bits (89), Expect = 9e-41
 Identities = 89/89 (100%)
 Strand = Plus / Minus

Query: 1614   agaaaaaagtccagtggaagcaaaattgccttggctgaaacaagcacaagaactagaaga 1673
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 116453 agaaaaaagtccagtggaagcaaaattgccttggctgaaacaagcacaagaactagaaga 116394

Query: 1674   gaccagaatagcaacagaagaaccaacgt 1702
              ||||||||||||||||||||||||||||
Sbjct: 116393 gaccagaatagcaacagaagaaccaacgt 116365
```

```
 Score =  149 bits (75), Expect = 2e-32
 Identities = 75/75 (100%)
 Strand = Plus / Minus

Query: 728    gagaagaaaatgtaattgctgttcctttgggaagtcgaagtgtgagaattacagtgaaag 787
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 157331 gagaagaaaatgtaattgctgttcctttgggaagtcgaagtgtgagaattacagtgaaag 157272

Query: 788    gacctgcccacctct 802
              |||||||||||||||
Sbjct: 157271 gacctgcccacctct 157257
```

```
 Score =  105 bits (53), Expect = 3e-19
 Identities = 53/53 (100%)
 Strand = Plus   Minus


Query: 1210   agtgatggatttaaagagataatgccctatgaccacttccaacctcttcctcg 1262
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 124771 agtgatggatttaaagagataatgccctatgaccacttccaacctcttcctcg 124719



>AC116157.3.1.75612
          Length = 75612

 Score =  547 bits (276), Expect = e-152
 Identities = 285/288 (98%)
 Strand = Plus / Minus


Query: 1701 gttcattccagaaccctggtcagcctgcagtaccacgtgtgggccaggtgtgcaggtccg 1760
          |||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct: 4268 gttcattccagaaccctggtcagcctgcagtaccacgtgtgggccgggtgtgcaggtccg 4209


Query: 1761 cgaggtgaagtgccgtgtgctcctcacattcacgcagactgagactgagctgcccgagga 1820
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4208 tgaggtgaagtgccgtgtgctcctcacattcacgcagactgagactgagctgcccgagga 4149


Query: 1821 agagtgtgaaggccccaagctgcccaccgaacggccctgcctcctggaagcatgtgatga 1880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4148 agagtgtgaaggccccaagctgcccaccgaacggccctgcctcctggaagcatgtgatga 4089


Query: 1881 gagcccggcctcccgagagctagacatccctctccctgaggacagtgagacgacttacga 1940
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 4088 gagcccggcctcccgagagctagacatccctctccctgaggacagtgagacgacttacga 4029


Query: 1941 ctgggagtacgctgggttcacccccttgcacagcaacatgcttgggagg 1988
          |||||||||||||||||||||||||||||||||||||| |||||||
Sbjct: 4028 ctgggagtacgctgggttcacccccttgcacagcaacatgcgtgggagg 3981



 Score =  408 bits (206), Expect = e-110
 Identities = 206/206 (100%)
 Strand = Plus / Minus


Query: 1262  gctgggaacataatccttggactgcatgttccgtgtcctgtggaggagggattcagagac 1321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 24689 gctgggaacataatccttggactgcatgttccgtgtcctgtggaggagggattcagagac 24630


Query: 1322  ggagctttgtgtgtgtagaggaatccatgcatggagagatattgcaggtggaagaatgga 1381
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct: 24629 ggagctttgtgtgtgtagaggaatccatgcatggagagatattgcaggtggaagaatgga 24570
```

```
Query: 1382  agtgcatgtacgcacccaaacccaaggttatgcaaacttgtaatctgtttgattgcccca 1441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 24569 agtgcatgtacgcacccaaacccaaggttatgcaaacttgtaatctgtttgattgcccca 24510


Query: 1442  agtggattgccatggagtggtctcag 1467
              |||||||||||||||||||||||||
Sbjct: 24509 agtggattgccatggagtggtctcag 24484


 Score =  305 bits (154), Expect = 1e-79
 Identities = 157/158 (99%)
 Strand = Plus / Minus


Query: 803   ttattgaatcaaaaacacttcaaggaagcaaaggagaacacagctttaacagccccggcg 862
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46577 ttattgaatcaaaaacacttcaaggaagcaaaggagaacacagctttaacagccccggcg 46518


Query: 863   tctttgtcgtagaaaacacaacagtggaatttcagaggggctccgagaggcaaactttta 922
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 46517 tctttctcgtagaaaacacaacagtggaatttcagaggggctccgagaggcaaactttta 46458


Query: 923   agattccaggacctctgatggctgatttcatcttcaag 960
              |||||||||||||||||||||||||||||||||||||
Sbjct: 46457 agattccaggacctctgatggctgatttcatcttcaag 46420


 Score =  297 bits (150), Expect = 4e-77
 Identities = 150/150 (100%)
 Strand = Plus / Minus


Query: 1466  agtgcacagtgacttgtggccgagggttacggtaccgggttgttctgtgtattaaccacc 1525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 19516 agtgcacagtgacttgtggccgagggttacggtaccgggttgttctgtgtattaaccacc 19457


Query: 1526  gcggagagcatgttgggggctgcaatccacaactgaagttacacatcaaagaagaatgtg 1585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 19456 gcggagagcatgttgggggctgcaatccacaactgaagttacacatcaaagaagaatgtg 19397


Query: 1586  tcattcccatcccgtgttataaaccaaaag 1615
              |||||||||||||||||||||||||||||
Sbjct: 19396 tcattcccatcccgtgttataaaccaaaag 19367
```

Score = 280 bits (141), Expect = 8e-72
Identities = 141/141 (100%)
Strand = Plus / Minus

```
Query: 1071  aggttatcagctcaattctgctgaatgtgtggatatccgcttgaagagggtagttcctga 1130
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27265 aggttatcagctcaattctgctgaatgtgtggatatccgcttgaagagggtagttcctga 27206


Query: 1131  ccattattgtcactactaccctgaaaatgtaaaaccaaaaccaaaaactgaaggaatgcag 1190
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 27205 ccattattgtcactactaccctgaaaatgtaaaaccaaaaccaaaaactgaaggaatgcag 27146


Query: 1191  catggatccctgcccatcaag 1211
              ||||||||||||||||||||
Sbjct: 27145 catggatccctgcccatcaag 27125
```

Score = 230 bits (116), Expect = 7e-57
Identities = 116/116 (100%)
Strand = Plus / Minus

```
Query: 959   agaccaggtacactgcagccaaagacagcgtggttcagttcttcttttaccagcccatca 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 32273 agaccaggtacactgcagccaaagacagcgtggttcagttcttcttttaccagcccatca 32214


Query: 1019  gtcatcagtggagacaaactgacttctttccctgcactgtgacgtgtggaggaggt 1074
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 32213 gtcatcagtggagacaaactgacttctttccctgcactgtgacgtgtggaggaggt 32158
```

Score = 176 bits (89), Expect = 9e-41
Identities = 89/89 (100%)
Strand = Plus / Minus

```
Query: 1614  agaaaaaagtccagtggaagcaaaattgccttggctgaaacaagcacaagaactagaaga 1673
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 17727 agaaaaaagtccagtggaagcaaaattgccttggctgaaacaagcacaagaactagaaga 17668


Query: 1674  gaccagaatagcaacagaagaaccaacgt 1702
              |||||||||||||||||||||||||||||
Sbjct: 17667 gaccagaatagcaacagaagaaccaacgt 17639
```

```
 Score =  149 bits (75), Expect = 2e-32
 Identities = 75/75 (100%)
 Strand = Plus / Minus


Query: 728   gagaagaaaatgtaattgctgttcctttgggaagtcgaagtgtgagaattacagtgaaag 787
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 58620 gagaagaaaatgtaattgctgttcctttgggaagtcgaagtgtgagaattacagtgaaag 58561


Query: 788   gacctgcccacctct 802
             |||||||||||||||
Sbjct: 58560 gacctgcccacctct 58546


 Score =  105 bits (53), Expect = 3e-19
 Identities = 53/53 (100%)
 Strand = Plus / Minus


Query: 1210  agtgatggatttaaagagataatgccctatgaccacttccaacctcttcctcg 1262
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26045 agtgatggatttaaagagataatgccctatgaccacttccaacctcttcctcg 25993


>AC116157.3.94489.171225
          Length = 76737

 Score =  472 bits (238), Expect = e-129
 Identities = 238/238 (100%)
 Strand = Plus / Minus


Query: 364  gactgccctccagatgcagaagatttcagagcccagcagtgctcagcctacaatgatgtc 423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3187 gactgccctccagatgcagaagatttcagagcccagcagtgctcagcctacaatgatgtc 3128


Query: 424  cagtatcaggggcattactatgaatggcttccacgatataatgatcctgctgccccgtgt 483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3127 cagtatcaggggcattactatgaatggcttccacgatataatgatcctgctgccccgtgt 3068


Query: 484  gcactcaagtgtcatgcacaaggacaaaacttggtggtggagctggcacctaaggtactg 543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3067 gcactcaagtgtcatgcacaaggacaaaacttggtggtggagctggcacctaaggtactg 3008


Query: 544  gatggaactcgttgcaacacggactccttggacatgtgtatcagtggcatctgtcagg 601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3007 gatggaactcgttgcaacacggactccttggacatgtgtatcagtggcatctgtcagg 2950
```

```
 Score =  260 bits (131), Expect = 8e-66
 Identities = 131/131 (100%)
 Strand = Plus / Minus


Query: 188   agacctcaagaaacactcgttcagatgaagacaaagatggcaactgggatgcttggggcg 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 49522 agacctcaagaaacactcgttcagatgaagacaaagatggcaactgggatgcttggggcg 49463


Query: 248   actggagtgactgctcccggacctgtgggggaggagcatcatattctctgcggagatgtt 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 49462 actggagtgactgctcccggacctgtgggggaggagcatcatattctctgcggagatgtt 49403


Query: 308   tgactggaagg 318
              ||||||||||
Sbjct: 49402 tgactggaagg 49392


 Score = 97.6 bits (49), Expect = 7e-17
 Identities = 49/49 (100%)
 Strand = Plus / Minus


Query: 316   aggaattgtgaagggcagaacattcggtacaagacatgcagcaatcatg 364
              |||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 18372 aggaattgtgaagggcagaacattcggtacaagacatgcagcaatcatg 18324


>AC116157.3.75713.94387
          Length = 18675

 Score =  260 bits (131), Expect = 8e-66
 Identities = 131/131 (100%)
 Strand = Plus / Minus


Query: 598  caggcagtgggctgcgatcggcaactgggaagcaatgccaaggaggacaactgtggagtc 657
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3634 caggcagtgggctgcgatcggcaactgggaagcaatgccaaggaggacaactgtggagtc 3575


Query: 658  tgtgccggcgatggctccacctgcaggcttgtacggggacaatcaaagtcacacgtttct 717
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 3574 tgtgccggcgatggctccacctgcaggcttgtacggggacaatcaaagtcacacgtttct 3515


Query: 718  cctgaaaaaag 728
             ||||||||||
Sbjct: 3514 cctgaaaaaag 3504
```

```
>AC087738.10.149851.181632
          Length = 31782

 Score =  244 bits (123), Expect = 5e-61
 Identities = 123/123 (100%)
 Strand = Plus / Plus


Query: 67    cagaccacagctgagaaatctcctggagcctatttccttcccgagtttgcactttctcct 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 9168  cagaccacagctgagaaatctcctggagcctatttccttcccgagtttgcactttctcct 9227


Query: 127   cagggaagttttctggaagacacaacaggggagcagttcctcacttatcgctatgatgac 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 9228  cagggaagttttctggaagacacaacaggggagcagttcctcacttatcgctatgatgac 9287


Query: 187   cag 189
              |||
Sbjct: 9288  cag 9290



>AC087738.10.1.149750
          Length = 149750

 Score =  137 bits (69), Expect = 8e-29
 Identities = 69/69 (100%)
 Strand = Plus / Plus


Query: 1       atggcttcctggacgagcccctggtgggtgctgatagggatggtcttcatgcactctccc 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 110273  atggcttcctggacgagcccctggtgggtgctgatagggatggtcttcatgcactctccc 110332


Query: 61      ctcccgcag 69
                |||||||||
Sbjct: 110333  ctcccgcag 110341
```